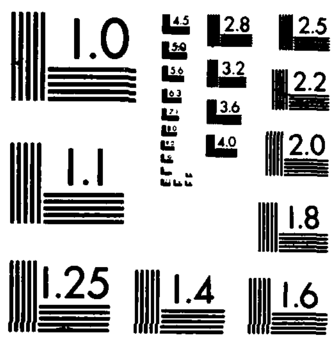
AD-A174 371 FEASIBILITY OF USING DATA ENVELOPMENT ANALYSIS AND REGRESSION ANALYSIS IN (U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF SYST J R WALLACE 1/2

UNCLASSIFIED SEP 86 AFIT/GSM/LSY/86S-21 F/G 5/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AF INSTITUTE OF TECHNOLOGY

FEASIBILITY OF USING DATA ENVELOPMENT
ANALYSIS AND REGRESSION ANALYSIS
IN COMBINATION TO INDICATE
EFFICIENT ALLOCATION OF
OPERATIONS AND MAINTENANCE FUNDS

THESIS

Jay R. Wallace, II
Captain, USAF

AFIT/GSM/LSY/86S-21

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GSM/LSY/86

FEASIBILITY OF USING DATA ENVELOPMENT ANALYSIS AND REGRESSION ANALYSIS IN COMBINATION TO INDICATE EFFICIENT ALLOCATION OF OPERATIONS AND MAINTENANCE FUNDS

THESIS

Jay R. Wallace, II
Captain, USAF

AFIT/GSM/LSY/86S-21

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information is contained therein. Furthermore, the views expressed in the document are those of the author and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the United States Air Force, or the Department of Defense.

AFIT/GSM/LSY/86S-21

FEASIBILITY OF USING DATA ENVELOPMENT ANALYSIS AND REGRESSION ANALYSIS IN COMBINATION TO INDICATE EFFICIENT ALLOCATION OF OPERATIONS AND MAINTENANCE FUNDS

THESIS

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Insistute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Systems Management

Jay R. Wallace, II, B.S.

Captain, USAF

September 1986

Approved for public release; distribution unlimited

Acknowledgements

I thank the Lord Jesus Christ for blessing me with the successful completion of this thesis. In the paragraphs that follow I recognize those He provided who were particularly helpful.

Major William Bowlin, my thesis advisor, was instrumental in distilling the thesis topic and invaluable in his reliably candid and beneficial feedback. I also thank Mr. Rich Murphy for his expertise and advice concerning regression analysis.

Of the many professionals at Headquarters Strategic Air Command who assisted me in gathering data, I especially thank Captain Jeff Steig, who served as my central point of contact, and Mr. Bert Berthelsen, whose Base Operations Support funds management branch offered frequent and enthusiastic assistance.

Finally, and with particular gratitude, I thank my lovely wife Patti for her continual encouragement and total typing support, despite the demands of our newborn and two-year-old daughters.

Jay R. Wallace, II

Table of Contents

List of Figures

List of Tables

Abstract

The present call for efficient use of national defense resources draws its urgency from the highest level of the American government. One problem with the Air Force's current method of allocating Operations and Maintenance ~~(O&M)~~ funds is that historical spending inefficiencies may get carried into the baseline for future funds allocations, thereby defeating this call for efficient use of resources.

This ~~research~~ furthers the exploration of ways to efficiently allocate resources and studies the feasibility of using a relatively new methodology developed by Charnes, Cooper, and Rhodes called Data Envelopment Analysis (DEA), in combination with regression analysis, as a method for indicating efficient allocation of Air Force O&M funds. Intended to illustrate O&M funds in general, this study uses a data base consisting of Fiscal Year ~~(FY)~~ 1985 Base Operations Support (BOS) obligations and activity measures from 25 Strategic Air Command ~~(SAC)~~ wings.

Research methodology consists of first, the derivation of a regression model expressing BOS obligations as a function of BOS activity measures; second, the development of an efficient data set by using information generated by DEA; and finally, the application of regression analysis to efficient data in order to predict

efficient BOS obligations for the SAC wings involved.

Research results indicate that if the sampled BOS obligations and activities had been at theoretically efficient levels, then the cumulative FY 1985 BOS obligations would have been $9.7 million less than actual. The study also concludes that these procedures show promise for use in establishing a funds allocation baseline based on efficient use of funding.

This research concludes with recommendations for further study into the areas of developing appropriate output measures, using indicator variables in the regression model, breaking down efficiency measures into their technical and scale efficiency components, and dealing with collinearity among efficient data.

I. Introduction

General Issue

The objective of efficiently using national defense resources draws its urgency from the highest level of the American government. In June 1982, President Reagan created the President's Private Sector Survey on Cost Control, popularly known as the Grace Commission (16:1). The Grace Commission, whose mandate was "to identify opportunities for increased efficiency and reduced costs achievable by executive or legislative action", submitted its final report to the President in January 1984 (16:2,3). Following a period wherein the General Accounting Office reviewed the Grace Commission's major proposals, the Comptroller General of the United States issued a summary report to the U.S. Senate Committee on Governmental Affairs. The executive summary of the Comptroller General's February 19, 1985 report includes the following statement:

> Reducing the huge federal budget deficit has become the most critical challenge facing the Congress and the President. To achieve any appreciable reduction in the deficit, a host of tough policy choices will have to be made regarding the level and composition of government services and the sources and distribution of the revenues necessary to finance those services. Moreover, those operations and programs that are continued will have to be conducted more efficiently and economically [27:i].

The call for efficient use of federal funding has

gone forth, obviously encompassing both the Department of Defense and the Air Force. As this chapter will later discuss, the Grace Commission is neither the first, nor the only acknowledgement of the importance of efficiently using scarce funding. Other events such as a joint Department of Defense (DOD)/General Accounting Office (GAO) study of the Planning, Programming, Budgeting System (PPBS) (28); a GAO report to Congress on budget resources, accomplishments, and problems (15); plus Air Force financial management regulations (such as AFRs 170-6 and 172-1 to be discussed later), all acknowledge the importance of efficiently using scarce funding.

In spite of the abundance of documented support for efficient use of funds, the need remains for a method to allocate Operations and Maintenance (O&M) funds in the Air Force in order to encourage efficient operations. Because O&M funding for the Air Force is a limited resource, budget operations should allocate the funds so they are used as efficiently as possible. The general issue of this research study, therefore, is to investigate how the Air Force can allocate its O&M funds so as to encourage efficient operations.

By way of overview, the remainder of this introductory chapter will address the following topics: first, a presentation of key definitions; second, a close look at events that have underscored the need for

efficient use of funds; third, a thorough problem statement of the need to efficiently allocate Air Force O&M funds, to include a survey of past and present initiatives to solve the problem; fourth, a discussion of this study's research objective, to include the specific research questions involved and the academic contribution intended; fifth, a discussion of the limitations and intended scope of this research; and finally, as a concluding note to this chapter, there will be a brief orientation to the remaining three chapters of this report.

Key Definitions

In resuming discussion of the general issue of efficient use of federal funding, the first order of business is to establish a common understanding of three key terms: productivity, effectiveness, and efficiency. Looking first to productivity, according to Cooper and Ijiri, productivity is:

> The yield obtained from any process or product by employing one or more factors of production. Productivity is usually calculated as an index number; the ratio of output to input [17:399].

Additionally, Mali defines productivity as:

> the measure of how well resources are brought together in organizations and utilized for accomplishing a set of results. Productivity is reaching the highest levels of performance with the least expenditure of resources [14].

Mali also offers a more condensed definition of productivity stating that "productivity is a combination of

effectiveness and efficiency" (14). To understand Mali's second definition, however, effectiveness and efficiency need to be defined.

Effectiveness, according to Cooper and Ijiri, is the "Ability to (a) state and (b) achieve objectives" (17:190). According to Webster's New College Dictionary, effectiveness is defined as "producing a decided, decisive or desired effect, or being capable of producing a desired effect" (14). And finally, an even more simple definition, this one coming from DOD Instruction 5010.34 "Productivity Enhancement, Measurement, and Evaluation - Operating Guidelines and Reporting Instructions", states that "effectiveness means accomplishing the right things in the right quantities, at the right times" (14).

The third term to be defined here, and the most important to this research study, is "efficiency". Cooper and Ijiri generally define efficiency as "A conventional measure of performance expressed in terms of a standard of comparison; applied to a machine, operation, individual, or organization"; however, they also specifically label efficiency as "The ratio of output to input (17:191). Maintaining the focus on output and input, Charnes, Cooper, and Rhodes define efficiency by considering the concept from two different orientations:

> Output Orientation: An organization is not efficient if it is possible to increase any output without increasing any input or decreasing any other output.

Input Orientation: An organization is not efficient if it is possible to decrease any input without increasing any other input and without decreasing any output [14].

Although the next chapter of this report will further define efficiency in terms of specific types--such as relative efficiency, technical efficiency, and scale efficiency--the basic definitions of productivity, effectiveness, and efficiency provided above should suffice for the introductory discussion that immediately follows.

The Need for Efficient Use of Funds

As mentioned at the outset of this chapter, the Grace Commission's assigned objective was basically to identify ways to improve efficiency and cut costs in the federal government. The end result of the Grace Commission's study took the form of 47 seperate reports, addressing 784 different issues, and recommending 2,478 specific actions, which if implemented were estimated to result in "net savings and revenue increases of $424.4 billion over three years" (16:3). Looking specifically under the category of national defense, the Grace Commission offered 112 recommendations equating to an estimated $94 billion savings (16:27).

Interestingly enough, several actions recommended by the Grace Commission were already being pursued by DOD as a result of the DOD Acquisition Improvement Program, popularly known as the Carlucci Initiatives.

Having its inception in April 1981, the Carlucci Initiatives began a major series of acquisition related reforms including multiyear procurement, dual-sourcing for procurement, and more efficient production rates--three actions which were also major recommendations by the Grace Commission. In fact, 33 of the 112 Grace Commission recommendations directed at the DOD were similar to actions already being taken as a result of the earlier Carlucci Initiatives (16:33).

Another recent event underscoring the need for efficient use of funding was the 1983 Joint DOD/GAO Working Group on PPBS(28). Briefly defined, "PPBS is the framework for the resource allocation decision process that is driven by the plans, programs, and budget decisions made by DOD management, the President, and the Congress" (28:15). The Working Group, in an effort to identify potential improvements for the DOD's formal resource allocation process, issued a report describing possible improvements in each of nine selected areas of study. The study area discussed in Section 7 of the Working Group's report was titled "Programmatic Analysis of Operating Accounts" and was defined as follows:

> This study area concerns the operating accounts--namely the operations and maintenance and military personnel accounts--and how they are analyzed in PPBS program development for establishing output objectives and the required resource inputs. More specifically, the area concerns the steps that are

> being taken, or remain to be taken, to improve procedures for analyzing how various resource levels and mixes in the operating accounts can affect our armed services' capability for waging war [28:91].

One of the reasons why the above study area was chosen was that DOD managers, civilian and military alike, knew that with an increase or decrease in efficiently used funding there should be a commensurate increase or decrease in fighting capability--yet no such resource-to-readiness analysis was available (28:91). Though the Working Group acknowledged that since the late 1970s DOD had "significantly intensified their efforts to study and report on readiness and sustainability and related funding questions", they also reported the potential for "further improvements in efforts to analyze and identify the effects of alternative funding levels and mixes in operating accounts upon readiness and sustainability levels" (28:91,92). The Working Group's specific recommendation concerning analysis of input-to-output relationships in operating accounts is shown below.

> OSD could initiate and coordinate OSD-JCS-military department efforts to develop more comprehensive analyses of the relationship between different operating account resource inputs and the effect on aggregate readiness and sustainability goals [28:96].

Stated differently, in order to know what increase in output (in this case readiness and sustainability) should be expected from an increase in input (operating

account dollars), the efficiency of the process converting input into output must be known. With such measures of efficiency available for different processes (i.e., organizations), a number of benefits become available. First, the less efficient organizations could make improvements based on knowledge gained from the more efficient organizations. Second, scarce funding could be allocated based to some extent on the efficiency of its recipients. And third, increments and decrements of funding could perhaps be accompanied with a predominantly objective measure of impact on output.

In step with recommendations of the Grace Commission and the DOD/GAO Working Group, the Air Force is by no means unaware of the imperative for efficient use of limited funding. Though the following discussion includes only a sampling of three Air Force financial management related publications, the theme of efficiency awareness is clear.

Air Force Regulation (AFR) 170-6 "Comptroller Activities, Functions, and Responsibilities" defines the overall objective of the Air Force Comptroller to include:

> establishing responsible and professional financial management throughout the Air Force to promote financial integrity, outstanding analytical capabilities, and economical and efficient use of resources [18:3].

In fact, one of the Comptroller's specific functional responsibilities is "achieving economies and efficiencies

in the use of financial resources and limiting fraud, waste, or abuse in the use of these resources" (18:2).

A second regulation, AFR 172-1 "USAF Budget Manual: Budget Management for Operations", speaks particularly about the Economies and Efficiencies (E&E) Program stating its intent is "to create a financial and management environment that encourages and rewards improved efficiency" (19:74). The Air Force's seven formal E&E related programs are listed below (19:74).

1. Productivity Investment Program
2. Fast Payback Capital Investment Program
3. Value Engineering
4. Air Force Suggestion Program
5. Air Force Management Engineering Program
6. Energy Conservation Investment Program
7. Air Force Productivity Enhancement Program

Whether benefits from the above programs accrue to the O&M appropriation or to various investment appropriations, the common theme is that of getting the most out of every budget dollar.

The third and final reference draws from Air Force Pamphlet 172-4 "The Air Force Budget Process." As the pamphlet is quoted below, resource management--right down to the base level--must concern itself with the need for wise use of funding.

> Managers at base level are keenly aware of the shortage of resources available for

> achieving goals and the difficult choices that this limitation imposes. Some goals must be deferred, and some programs eliminated or reduced in light of the austere funding levels in which managers must continue to operate [20:71].

With the preceding references to the Grace Commission, the Joint DOD/GAO Working Group on PPBS, and selected Air Force publications serving to substantiate the general issue of efficient use of funding, the next portion of this introductory chapter will focus on the more specific problem of allocating O&M funding to encourage its efficient use.

Problem Statement

Basically, the current method for allocating O&M funds involves distribution based on an entity's historical spending, plus or minus amounts corresponding to changes in that entity's mission. The problem with the current method is that historical spending inefficiencies may get carried into the baseline for future funding allocation. According to Clark (14), when budgeting for similar units is based on historical consumption "an inefficient unit which uses excessive amounts of resources may receive more money than an efficient unit." Bowlin, in further agreement with the problem of efficient funds allocation, makes the following statement:

> If an Air Force wing has been operating inefficiently in the past, historical costs will include the cost of operating inefficiently. Using historical costs as a factor

> in distributing funds and other resources may reward inefficient wings by providing funds and other resources which serve to perpetuate and perhaps increase inefficiencies [7:8].

To avoid perpetuating such inefficiencies, budget operators need a management tool indicating what the efficient allocation of O&M funds should be for an entity, based on that entity's characteristics.

In the author's experience as a budget officer at both the wing and major command levels, organizations--regardless of their efficiency--tend to fully obligate the O&M funds allocated to them. The situation behind an organization's total obligation of funds can certainly vary: some organizations may be underfunded, running out of funds before they run out of mission; some may be properly funded, having just the amount of funding necessary to accomplish the mission efficiently; some may be overfunded, taking advantage of the opportunity to satisfy previously deferred unfunded requirements; and some may be avoidably inefficient, concentrating on spending what they have this year for fear of losing it next year. Whichever situation applies, efficient allocation of funding would offer a framework of incentive for the less efficient organizations to learn from the more efficient organizations. Holding input constant, the only way the Air Force can get more out of its budget dollar is to reduce inefficient spending in the process of converting funding input into mission output.

Observers, both inside and outside DOD, are keenly aware of how important it is to efficiently allocate scarce national defense operating dollars. Ms. Karen Alderman, who in 1983 was the Acting Deputy Assistant Secretary of Defense for Civilian Personnel Policy and Requirements in the Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs and Logistics), says concerning efficiency:

> The economic conditions facing the U.S. today mandate that the Defense Department manage its funds as prudently and efficiently as possible. Managers must continually question whether everything that they are doing is really necessary to accomplish the missions [1:17].

In the same article, Alderman points out that by conducting efficiency reviews and discovering more economical ways to do DOD missions, efficiency improvements can equate to real payoffs relative to fixed funding levels (1:17).

In agreement with Alderman, Mr. Vincent Byrne, speaking to the American Society of Military Comptrollers as Director of Administration for Xerox Corporation, encourages his audience to solve their operational shortfalls by increasing productivity of current resources, not by anticipating future increases in resource levels. Byrne specifically underlines the need for efficient operations by emphasizing the need for "fuller use of available resources and current funding levels" (9:7).

Lastly, with particular attention to DOD O&M funding, the General Accounting Office (GAO) in a 1983 report on the defense budget states the following:

> DOD did not have a well-planned strategy and priority system for applying increased funding to O&M programs. As a result, funds were applied to some programs in excess of what could be absorbed efficiently and effectively [15:iii].

In the same report, the GAO specifically critiques DOD's technique for accountability of program execution. The report states:

> The budget execution goals and objectives established for the programs we reviewed are based on consumption requirements, such as how much of the appropriations have been obligated. Instead, resources should be related to achievements and how much the additional expenditures are expected to increase capability [15:iv].

Finally, speaking specifically about Real Property Maintenance and Repair programs, a subaccount of O&M, the GAO report states:

> We found no accountability systems linking military capability and rising or falling program funding levels. Budget estimates are often limited to prior-year funding plus anticipated program growth and inflation. Since funding is not linked to intermediate outputs, such as increased proficiency or mission capable weapon systems, or to ultimate outputs, such as increased readiness, there is no way of determining if the services could achieve the same goals with fewer dollars [15:24].

As expressed above by Alderman and Byrne, and particularly by the GAO, efficient allocation of the national defense operations dollar is a must.

Survey of Previous Initiatives. In this discussion of the problem surrounding efficient allocation of Air Force O&M funding, it is important to realize that it is neither a new problem nor has it gone without prior attempts to solve it. In fact, this research effort is only the most recent in a lengthy continuum of effort to efficiently allocate defense dollars.

Back in 1962, under Secretary of Defense McNamara, the Air Force implemented "program budgeting", a concept first presented by David Novick in a 1954 Rand Corporation report (7:22). Program budgeting--which eventually evolved into the present day Planning, Programming, Budgeting System (PPBS)--was generally intended to "provide a more efficient and effective allocation of resources" and to "improve Defense Department and Air Force ability to make rational choices among alternative uses of funds" (7:22,23). More specifically though, Bowlin lists four things which program budgeting was supposed to accomplish in the Air Force:

1. Focus attention on the need for explicit definition of the objectives by reference to expected outputs as well as the inputs required to sustain various Air Force programs.

2. Facilitate systematic analyses of possible alternative programs by reference to trade-off possibilities for meeting those objectives.

3. Provide a basis for comparative (quantitative) evaluation of the benefits and

costs of different programs which in military terms, takes the form of cost and effectiveness measures.

4. Produce total rather than partial cost estimates of programs across relevant periods of time [7:24].

Though PPBS is still DOD's formal process for making resource allocation decisions, the 1983 Joint DOD/GAO Working Group on PPBS (discussed earlier) has offered recommendations to improve PPBS's weakness in analyzing its operating accounts.

Continuing the effort toward efficient allocation of resources, in the late 1960's the Air Force applied the Output Measurement Program (OMP) and the accompanying Cost Center Performance Measurement System (CCPMS). Aimed at increasing the emphasis on efficiency in budgeting, the programs were to "help managers identify problem areas regarding the efficient use of resources and thereby improve resource allocation and management by correlating output with cost" (7:24,25). Unfortunately, due to the following three reasons cited by Covell and Jones, the OMP and CCPMS programs met with failure and eventual cancellation in 1977 (7:25,26).

1. The output measures used by the programs inadequately described their respective cost center's activities.

2. The measures of inefficiency fed back to managers were not specific enough to facilitate problem identification and corrective action.

3. The regression model used in CCPMS was not

helpful in predicting the resource impact resulting from a change in mission.

The most recent major program targeting efficient allocation of resources was Zero Based Budgeting (ZBB). ZBB, made part of the federal budget process by President Carter in 1979, was designed to give managers the information they needed to make sound decisions about the "efficient and effective use of resources" (7:26,27). Peter Phyrr, developer of ZBB, describes his budgeting system as posing two basic questions to organizations and their management. First, "Are the current activities efficient and effective?", and second, "Should current activities be eliminated or reduced to fund higher-priority new programs or to reduce the current budget?" (7:27). Though ZBB was formally discontinued because of its cumbersome nature, the influence of its "decision package" orientation remains evident in the priority ranking system of current Air Force budgets.

Before departing from the historical perspective of efficiency improvement efforts, a concluding post script about the ill-fated Cost Center Performance Measurement System (CCPMS) is warranted. P.S. Lein (a lifelong professional in the comptroller career field), referring among other things to the failure of CCPMS, offers the following advice:

> [D]on't assume that simply because an idea has surfaced before and has been discarded, that it

> was necessarily a bad idea. It may have been mis-applied; it may have simply had bad press; it may have been before its time; it may need only some minor adjustments to make it work [23:30].

Fortunately, as the rest of this paper indicates, the failure of CCPMS was *not* an epitaph for the idea of efficiently linking funding inputs with mission outputs.

Recent research studies have shown there are indeed analytical techniques available which make it feasible to rationally associate the allocation of funding with measures of efficiency. Bowlin, for instance, identified the possibility of applying Data Envelopment Analysis (DEA) to resource allocation (7:161). Bowlin proposed that DEA, a mathematical model for measuring organizational efficiency, could be combined with goal programming, a linear programming application which seeks an optimum solution while minimizing deviations from goal constraints. Bowlin asserted that such a combination would allow for an optimum allocation of funds across organizations, while satisfying the criteria of efficient goal constraints derived from DEA (7:161). While the theoretical underpinnings needed for this effort have been provided by Charnes, Cooper, Golany, Seiford, and Stutz (10), their theory has yet to be operationalized and tested in application to actual data. Thus, there is additional research to be accomplished before Bowlin's proposal can be fully implemented.

Montemayor, another recent researcher of efficient resource allocation, had the objective of developing "specific management techniques for dealing with the resource allocation problem through the use of Constrained Facet Analysis and the management information that it produces" (25:4). More specifically, Montemayor investigated the feasibility of using efficiency measures from Constrained Facet Analysis--a variant of DEA--for developing parameters for a general networking resource allocation model (25:ii). Bowlin, however, characterizes Constrained Facet Analysis as a method of efficiency measurement whose estimates are erratic, challengeable, and in need of further development (7:63-66).

As evidenced in discussion to this point, various reports and studies have shown a continuing need for determining a way to allocate funding in order to encourage efficient operations. Although Bowlin's and Montemayor's research have provided insights and progress into modeling efficient resource allocation, limitations in their research have left the door open for further exploration of the subject area.

Research Objective

As the following section of this chapter will discuss, the present research endeavors to join Bowlin and Montemayor in their exploration of ways to efficiently allocate resources. Recollecting earlier discussion

about the call which has gone forth for efficient use of federal funding, the reason for this research is to respond to that call by studying the feasibility of a technique intended to facilitate development of an efficient allocation of O&M funds. As trusted stewards of tax dollars intended for operating and maintaining national defense, "budgeteers" stand to benefit from sophisticated quantitative techniques which better equip them to offer commanders sound advice on efficient allocation of scarce funding. The specific objective of this research study, therefore, is to determine the feasibility of using Data Envelopment Analysis (DEA) in combination with regression analysis to provide a method for indicating an allocation of Air Force O&M funds for efficient operations.

The reasons for choosing to combine DEA and regression analysis to indicate efficient resource allocation are rather straightforward. DEA, a relatively recently developed mathematical model for measuring organizational efficiency, enumerates efficiency ratings and respective amounts of input surpluses and output shortages. Armed with such knowledge of inefficiencies and their sources, actual input and output data can be adjusted (via filling shortages and subtracting surpluses) to represent theoretically efficient values. At this point, the familiar process of regression analysis can be used to fit (via

least squares estimation) a regression function explaining how efficient resource inputs should vary with efficient organizational output levels.

Very briefly, the approach for accomplishing this research study will involve the following three steps: first, the identification and selection of organizational resource inputs and outputs (via expert opinion, a priori judgement, and regression analysis) to determine how level-of-output predicts input-of-resources; second, the application of DEA to the original input and output data to determine what the efficient data values would be; and third, the application of regression analysis to the efficiency-adjusted data and subsequent comparison of regression results based on original versus efficiency-adjusted data.

Guiding the approach described above, the primary research question of this study addresses application of DEA and regression analysis to Base Operations Support (BOS) funding of selected Strategic Air Command bases to indicate efficient allocation of BOS funds to those bases. Other research questions, preparatory to addressing the primary research question follow:

1. What base-level organizations receive their funding from the BOS program element?

2. Of all BOS-funded organizations, which receive the greatest share of BOS funding?

3. For those organizations receiving the greatest

share of BOS funding, what are the quantitative measures for their respective outputs?

4. Of those output measures, which are the best "cost drivers" or predictors of a base's total annual BOS dollar obligations?

5. What are the differences between each base's actual funding input and mission output data when compared to input and output data that has been adjusted to reflect efficiency?

6. How do regression analysis results differ between using actual base input and output data, as opposed to using data adjusted to reflect efficiency?

In terms of the contribution offered by this research, the benefit is twofold. First, from an academic perspective, this research explores the applicability of Data Envelopment Analysis to resource allocation by using its efficiency information for regression analysis. And second, from a practical perspective, it continues the advancing effort to achieve a method for efficient allocation of scarce funding.

Scope and Limitations

As stated in the title of this report, the thrust of this research study is one of exploring feasibility. Within its controlled scope, this study investigates the basic concept of predicting efficient funding inputs based on known mission requirements or outputs. To achieve the goal of discerning efficient input/output relationships, this research is limited to involve only Strategic Air Command bases, only the BOS program element,

and only primary cost drivers within that program element. Consequently, by employing judgement in place of random sampling for the selecting of data, the generalizability of findings is disallowed in the strict statistical sense. However, by intentionally designing this research as a prototypic study--one small enough to be relatively easily understood, yet large enough to reasonably represent the real world--the conclusions offered should give valuable insight into the task of efficiently allocating funds.

Research Development

The remaining three chapters of this thesis document the methodology, analysis, and conclusions of this research study. Chapter II, *Methodology*, describes the process of identifying and selecting inputs and outputs; introduces the concept of Data Envelopment Analysis; and explains how DEA and regression analysis are combined. Chapter III, Analysis, reports the results of the input and output identification and selection process; the efficiency information provided by application of DEA to input and output data; and, a comparison of regression results based on original versus efficient data. Finally, Chapter IV, Conclusions and Recommendations, summarizes the research findings and offers possible courses of further research.

II. Methodology

Overview

This study's primary research question addresses the feasibility of combining DEA and regression analysis to indicate an efficient allocation of Base Operations Support (BOS) funding. This chapter describes the three-step methodology involved in that effort.

The first step involves selecting the decision making units (organizations) for inclusion in the study; identifying quantifiable BOS inputs (funding) and outputs (activity); and finally, selecting by expert opinion, *a priori* judgement, and regression those outputs which best explain or predict variation in input. The result of this step is the identification of appropriate input and output variables for use in this study and the development of a regression model wherein significant BOS outputs are the independent variables, and BOS input is the dependent variable.

The second step involves using Data Envelopment Analysis, an efficiency identification model developed by Charnes, Cooper, and Rhodes (13), to identify sources and amounts of inefficiency in the selected data. Then, using this information and the procedures outlined by Charnes, Cooper, and Rhodes, a data set representing theoretically efficient inputs/outputs for all decision

making units (DMUs) is computed.

The third and final step takes the efficient input/output data from step two and develops a new regression model based on the efficient input/output data. In accordance with the objective of this research, the result of this final step is an efficiency-based regression model which theoretically predicts the efficient level of resource input necessary to accomplish given levels of BOS output. The predicted input values generated by the original and efficient regression models are then compared to assess the effect of using efficient data.

A thorough discussion of the above three steps follows, with greatest emphasis on a detailed discussion of Data Envelopment Analysis in step two. The data gathered and results of each methodology step are presented in Chapter III, Analysis.

Identifying and Selecting Inputs and Outputs

The first step of this research methodology involves the process of selecting the DMUs to be considered, identifying quantifiable BOS inputs and outputs for each DMU, and selecting those outputs which best explain or predict variation in input. The objective of this step is to identify appropriate input and output variables by developing a simple, yet reasonably valid regression model relating a DMU's level of funding input to that

DMU's level of activity output.

The author subjectively selected 25 Strategic Air Command (SAC) bases for study. The reason for selecting SAC bases it twofold--the author is familiar with SAC financial management practices and has many helpful personal contacts within HQ SAC. The base operations support (BOS) function at each SAC base is the decision making unit to be evaluated. BOS activity was selected because the composition of BOS activity is relatively similar among SAC bases. For instance, all SAC bases have standard BOS activities such as supply, transportation, and security police. (A thorough definition of SAC BOS activity is in Appendix A.) The reason for selecting a set of DMUs with similar activities (outputs) is to facilitate the Data Envelopment Analysis which takes place later in the study. Additionally, for the prototypic nature of this research, the input/output relationships within SAC's BOS program should adequately illustrate Air Force O&M funding in general.

Regarding identification and selection of inputs and outputs, the process involves the following steps. First, the author investigates the composition of SAC BOS activities by researching financial management regulations. Second, BOS funds management experts at HQ SAC are asked to identify what they consider to be the most significant components of BOS activity. Third,

experts in those most significant BOS activities are asked to provide quantifiable performance and activity measures. Fourth, in cases where such input and output measures are not available, the author attempts to identify logical factors which can be gathered from available BOS functional area data. Fifth, once all the candidate measures for BOS inputs and outputs are gathered, only those measures deemed appropriate are selected for use in this research.

Four guidelines for use in selecting appropriate input and output measures follow:

> First, the inputs and outputs should be comprehensive. That is, they should fully and properly measure...activity.
>
> Second, there should be some basis for believing that the relationships between inputs and outputs should be such that an increase in an input can reasonably be expected to increase one or more of the outputs.
>
> Third, all input and output measures should exist in positive amounts...
>
> Finally, the variables should be identifiable and defined and controlled so that they cannot be manipulated...[7:29,30].

Regarding the second guideline above, this research is more oriented to the perspective that a change in output level will result in a change in the required input level.

After using expert opinion and *a priori* judgement to select appropriate input and output measures, the sixth step involves applying regression analysis to the selected

inputs and outputs. By determining how well the measures of output (the independent variables) predict BOS obligations (the input or dependent variable), regression not only checks the findings of expert opinion and a priori judgement, but also provides added information regarding guideline number two, above. Ordinary least squares regression is used since it is a commonly used and well understood parametric technique for modeling the average relationship between independent variables and a dependent variable.

After completing the selection process by considering expert opinion, a priori judgement, and regression, the resultant regression equation, using BOS activity level to explain variation in BOS obligations, includes only the most significant output measures from the most significant BOS component organizations. The reason for limiting the number of output measures is to isolate the primary cost drivers behind BOS obligations, thus achieving the simplicity desired in this feasibility study. The overall purpose of developing this initial regression model is to later use its input and output variables in a second regression--one using efficient input and output data--thereby creating a new, efficient regression model.

Applying Data Envelopment Analysis (DEA)

The second step of this research involves using the original BOS input and output data (identified by the

selection process in the prior step) in a DEA, and thereby identifying sources and amounts of inefficiency in that data. The original input/output data of the DMUs which DEA identifies as inefficient are then adjusted to be efficient values as estimated by DEA. The reason for using DEA to generate efficient data is to permit a later regression operation, which will yield a model capable of indicating efficient BOS funding for a DMU, based on that DMU's BOS activity level.

Objective of DEA. Charnes, Cooper, and Rhodes, the developers of DEA, developed the procedure with the objective of measuring "relative efficiency" among decision making units (DMUs) (3:1370). A DMU is an organizational entity whose management makes decisions about using input resources in the production of outputs. The following is Charnes' definition of relative efficiency:

> 100% relative efficiency is attained by any DMU only when comparisons with other relevant DMUs do not provide evidence of inefficiency in the use of any input or output [11:72].

As noted by Banker, the relative efficiency measure from DEA introduces an interesting twist to conventional mathematical programming. While mathematical programming is usually a planning tool to help managers evaluate alternatives prior to decision making, DEA reverses the relationship by measuring the relative efficiency of decisions after managers have made them (4:1078). Charnes (13:429) refers to such measurement as "decision making

efficiency," while Bessent (5:57) labels it "managerial efficiency." Simply stated, the idea of managerial efficiency is that when a manager makes a resource decision about an organization's inputs or outputs, the subsequently observed relative efficiency of the organization is an indicator of how good the decision was.

DEA's measure of relative efficiency is the product of two other measures of efficiency: technical efficiency and scale efficiency (3). Cooper and Ijiri define the two terms as follows:

> Technical efficiency: the ability to obtain the greatest possible output from a given input; or to produce a given output with the lowest possible amounts of input.
>
> [S]cale efficiency occurs at the optimum scale under the best mix of inputs and technical efficiency conditions [17:191].

As Charnes explains, DEA extends the normal single input to single output efficiency definition used in the natural sciences, to a model capable of handling multiple inputs and multiple outputs (11:63). The model provides not only a scalar measure of efficiency but also identifies sources and amounts of inefficiency within the multiple inputs and outputs. For instance, when the model identifies an inefficient input, that input may be at a surplus level, and therefore not contributing to output. Armed with such information, management can take corrective action to reduce or eliminate the inefficiency (11:64).

Further investigation of the model's objective reveals it was developed with special reference to not-for-profit organizations and government agencies (5:57). Banker confirms the developers' intended application of DEA by stating:

> The main uses for these ideas have been in evaluations of "management" and "program" efficiencies of decision making units (DMUs) of a not-for-profit variety such as schools, hospitals, etc. [4:1079].

Bessent, using DEA to determine comparative efficiency of public schools, sought to answer two questions. First, which schools were relatively efficient regarding their inputs and outputs? And second, in the relatively inefficient schools, which inputs were inefficiently used (5:58)? As was already mentioned, DEA is designed to answer both questions. Restating the model's objective, Charnes says, "Our measure is intended to evaluate the accomplishments, or resource conservation possibilities, for every DMU with the resources assigned to it" (13:443).

The final area of discussion involving the objective of DEA touches on the philosophy of properly applying the model. According to Bowlin, an important assumption prior to using DEA is whether or not it is possible for management to take corrective action against detected inefficiencies. If management is not in a position to change the present mix of inputs and outputs, then they might be wiser to apply regression analysis, which would predict

future performance based on averaging historical data within the organization's fixed characteristics (6:137). In the exerpt below, Charnes confirms Bowlin's philosophical distinction regarding the appropriate use of the model, but uses general terms.

> [R]egression might be used when general characterizations are of interest for purposes of policy analysis and prediction of future behavior of the entire ensemble of observations. DEA might be used when interest centers on individual observations and the institutions (=DMUs) to which they relate [11:61].

Since this research, however, is interested in investigating the basic feasibility of combining information from both regression analysis and DEA, it is not constrained by possible or probable difficulties in correcting inefficiencies in levels of input or output.

A second point of philosophy concerning DEA encourages managers not to let the model's use lapse into a mere classification system. Bowlin presents this point by emphasizing that the model's usefulness for managers far exceeds the relatively superficial results of classifying an organization as either relatively efficient or inefficient. For the manager, the capability to detect and analyze inefficiencies is definitely more important than the ability to rank order decision making units by their efficiency scores (6:132). In straightforward terms, Bowlin states:

> We should note that DEA results are not an end in themselves. They are better regarded as

> providing guidance to auditors or persons concerned with budgetary and program reviews in identifying the possible presence of inefficiencies [7:161].

Mechanics of DEA. This discussion advances further into the detail of DEA, centering on the actual process of measuring relative efficiency. A detailed presentation of the DEA model in both its fractional and linear programming forms is in Appendix B.

By describing the model's process in straightforward terms, Bessent offers a step-by-step look at how the DEA fractional model works. First, the user must identify the organizations or suborganizations that will comprise the decision making units (DMUs). Second, the user must identify the inputs and outputs applicable to the DMU. The selected inputs and outputs must be quantifiable. Third, the user places the inputs and outputs into the model. Fourth, the model puts the entered values into a ratio with an expression for outputs in the numerator, and an expression for inputs in the denominator. And fifth, the model compares the measured output of the DMUs with the measured input resources (5:59). Expanding on this fifth and pivotal step of the process, Bessent explains that the model's mathematical programming operates under two conditions:

> First, inputs and outputs are to be weighted in such a way that each unit is compared to all of the others in the set and is constrained not to be larger than the best input/output ratio observed for any DMU, and, second,

> weights are calculated to give the largest possible ratio value condition [5:59].

The "weights" Bessent refers to in the second condition are factors (v_i and u_r in Appendix B) computed by the model and used in determining the relative efficiency rating. (Charnes, Cooper, and Rhodes prefer to use the term transformation ratios instead of weights (13).) Restating the final step of the process and its conditions, Bessent writes, "In simple terms, all units are compared in order to locate the best ones in the set and to use these as the criterion of efficiency" (5:60).

Refocusing on the concept of relative efficiency, Charnes states, "All evaluations in a DEA analysis are effected by reference to subsets of DMUs which are rated as 100% efficient" (12:109). He further clarifies the model's evaluation process by stating:

> each DEA optimization tends to pick the reference sets for these evaluations from the available (efficient) DMUs which are most like the unit being evaluated [12:109].

Therefore, if the model evaluates a DMU as relatively inefficient, the model accompanies that evaluation with the identification of a reference set of relatively efficient DMUs. Because the efficient units in the reference subset are those most similar to the inefficient unit's production process, as represented by the ratio of outputs to inputs, this reference feature provides valuable information for follow-up analysis of the inefficient

unit (12:110).

To graphically illustrate what DEA does, Bowlin offers the example shown in Figure 1 (7:75-77). The X_1 and X_2 axes indicate respective amounts of two inputs used to produce one unit of output. Points A, B, C, D, and E represent five different DMUs, each using a different combination of inputs X_1 and X_2 to produce one unit of output. The solid line connecting points A, B, and C represents part of the isoquant for one unit of output. In the vernacular of DEA, points A, B, and C represent DMUs with relative efficiency measures (h_0 in Appendix B) of 1.0 or 100%. Points A, B, and C are all efficient because "There is no point that can be generated from convex combinations of members of the production possibility set that will dominate them" (7:77). Points D and E, however, are inefficient since they are respectively dominated (enveloped) by D' and E'. Though D' and E' are not actual observations, they are convex combinations of A and B and B and C, respectively, and thus represent "elements of the efficient frontier production possibility set" (7:77). DEA's efficiency measures for the inefficient DMUs at points D and E, therefore, equate to the following distance ratios: $\overline{OD'}/\overline{OD}$ for point D and $\overline{OE'}/\overline{OE}$ for point E. The values of both ratios are clearly less than unity, thus indicating their relative inefficiency. Finally it should be noted that

points A and B form the efficient reference subset for point D, whereas points B and C form the reference subset for point E.

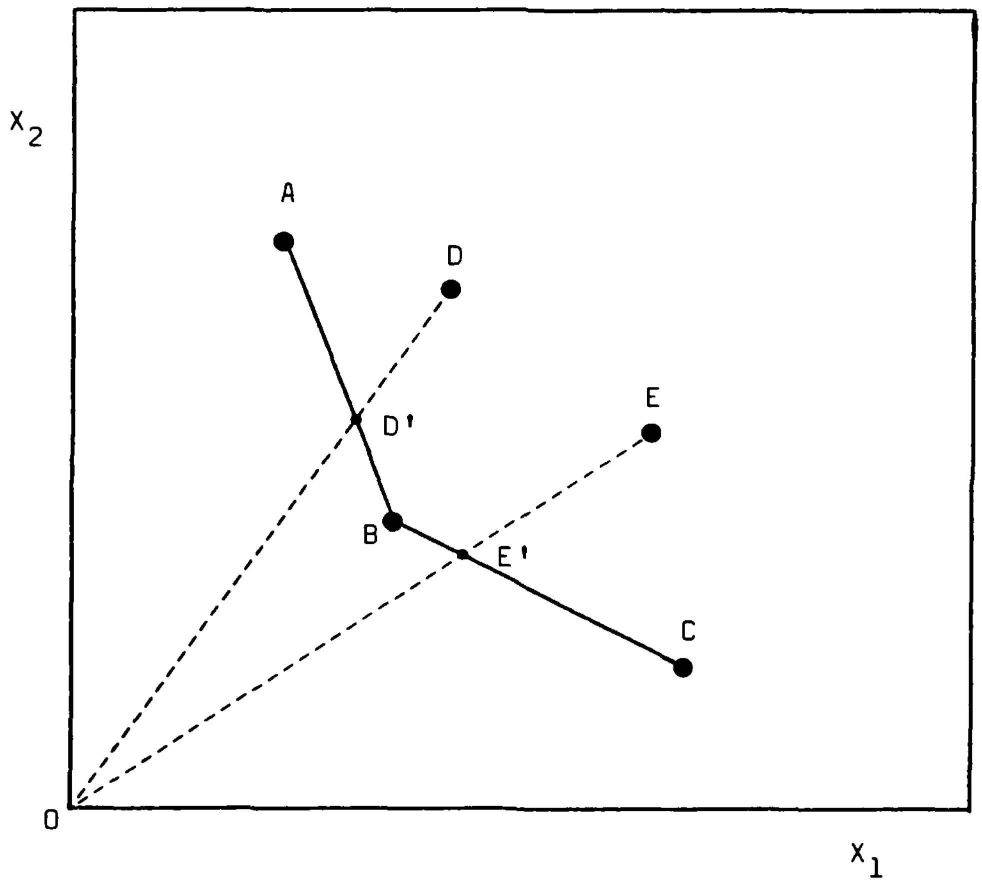


Figure 1. Example of Data Envelopment Analysis

As an interim summary of the DEA model, the following is a list of its primary characteristics (14):

1. DEA evaluates multiple outputs and multiple inputs simultaneously.

2. DEA handles non-commensurate units of measure among the outputs and inputs.

3. DEA identifies the sources and amounts of a DMU's inefficiencies in terms of output shortage and input surplus.

4. DEA computes respective output and input weights deterministically, rather than assigning weights _a priori_ via a preconceived production function.

5. DEA determines optimal output and input levels by using the extremal methodology of evaluating all DMUs relative to DMUs that are 100% efficient--not relative to some statistically averaged reference point which reflects both efficient and inefficient data.

Because of the advantages represented by the above list of characteristics, DEA is particularly well suited for use in the present research into efficient funds allocation. For instance, since there is no theoretical maximum level of output per dollar of BOS funding input, any related measures of efficiency must be of a relative rather than an absolute nature. Also, in the grand scheme of desiring to eliminate inefficient allocation of funds, the sources and amounts of inefficiencies (as identified by DEA) must be detected before they can be corrected.

DEA is not without its limitations, however. As described below by Banker, Charnes, Cooper, and Maindiratta (2), there are situations where DEA may not

properly determine an observation's relative efficiency.

> [T]he observations that are likely to be misclassified as efficient by the DEA model are "corner" points with a very small or very large quantity for at least one of the inputs or outputs, because of the inadequacy of referent points in the "corners" of a production possibility set [2:22].

Validation and Application of DEA. Recalling the earlier mention of regression analysis as a tool for predicting future performance based on historical data and fixed circumstances, Charnes strongly endorses DEA over regression analysis stating "DEA provides a better fit to each observation and a better basis for identifying and estimating the sources of inefficiency" (11:61).

In related research, as a result of investigating the accuracy and comparative merit of DEA and regression in estimating efficient input and output values, Bowlin, Charnes, Cooper, and Sherman (6) offer strong validation of DEA results. In their study, DEA and regression are individually applied to a contrived data set of 15 hypothetical hospital DMUs--some having efficient levels of input and output and others having inefficient levels. With the amounts and sources of inefficiencies known to the researchers, they were able to compare the abilities of DEA and regression to identify the inefficient DMUs and estimate the efficient input and output values. Bowlin, Charnes, Cooper, and Sherman (6:135) conclude that "the DEA estimates are almost uniformly better than

even the highly creditable regression estimates."

In another supportive study, this one comparing the performance of DEA and a parametric translog model in estimating a known production frontier, Banker, Charnes, Cooper, and Maindiratta conclude:

> The DEA models performed better than the translog model in estimating technical and scale efficiencies, rates of substitution, returns to scale and most productive scale size from the set of randomly generated observations [2:i].

Finally, Bowlin (6:136) further contends that the DEA model is superior not only to regression analysis, but also to ratio analysis (a single output to single input approach). Bowlin's hearty endorsement of DEA, however, does not imply he recommends exclusive use of the model. On the contrary, Bowlin endorses using various combinations of DEA, regression analysis, and ratio analysis (6:138). It is precisely this encouragement which leads this research to investigate the feasibility of using DEA and regression analysis, in concert, to predict efficient allocation of O&M funds.

Prior to this research, specific areas in the Air Force where researchers have applied DEA's objective of measuring relative efficiency include: real property maintenance activities (7), Tactical Air Command aircraft maintenance operations (12), Air Force Logistics Command depot-level maintenance (22) (24), and fire departments (8). In general, prior research applications of DEA

have highlighted the importance (and sometimes difficulty) of securing appropriate input and output measures; plus the high-value potential of management information about sources and amounts of inefficiency.

Of particular interest for this research is the earlier mentioned study by Montemayor, wherein he investigated the feasibility of using efficiency measures from Constrained Facet Analysis--a variant of DEA--for developing parameters for a general networking resource allocation model (25:ii). Montemayor's basic finding was that by efficiently reallocating the two inputs and two outputs of his study's 12 ficticious tactical fighter wings, higher levels of productivity became possible for the originally inefficient organizations (25:67,68).

Summarizing this discussion of the application of DEA, the model has been discussed in terms of its objective, its process, its mathematical form (Appendix B), its validation, and its application. The ultimate outcome of applying DEA in this research is a data set adjusted to indicate efficient BOS input and output values for each SAC wing. As discussed in the next section of methodology, this efficient data is used to develop an efficiency-based regression model capable of indicating efficient allocation of BOS funds, based on BOS activity levels.

Regression of Efficient Data

Using the same independent and dependent variables arrived at by the earlier regression process, this final step in the methodology involves a second application of regression--this time using the efficient input and output values generated by DEA. Based on the differences between original and adjusted data, the predicted input values of the second regression equation (based on efficient data) should differ from the predicted input values of the first regression equation (based on original data). The purpose of this step, and indeed the ultimate thrust of this research, is to generate a model capable of predicting an efficient allocation of BOS resources.

As mentioned earlier, because the focus of this research is on feasibility, it is not constrained by the aspect of how much discretion managers have to correct inefficient input and output levels. From a practical perspective then, the regression model generated by efficient data assumes all detected inefficiencies can be corrected. By comparing the predicted input values of the original and efficient regression models, therefore, the maximum effect of using efficient data is assessed. It is at this juncture where the contribution of DEA to the task of efficient allocation of BOS funding becomes evident.

Methodology Summary

To review, this study's methodology consists of three steps. The first step identifies and selects key inputs and outputs of SAC BOS activities, and from them generates a regression model predicting BOS funding input as a function of BOS activity level. The second step applies DEA to the original input/output data in step one, thereby identifying sources and amounts of inefficiency in each DMU. Detected inefficiencies in input and output values are then corrected, giving rise to an efficient data set. Finally, the third step performs a second regression analysis using the input/output variables from step one and the efficient data from step two, thereby generating a *new model* capable of efficiently allocating BOS funding based on assigned levels of BOS activity.

The methodology described in this chapter has two somewhat unusual aspects. First, regression analysis usually uses actual historical observations to develop a model for predicting future outcomes. In this study, regression analysis uses efficiency-adjusted data to build a model predicting theoretically efficient outcomes. The other unusual aspect involves the assignment of inputs and outputs. In this study, the DEA input is a wing's actual BOS dollar obligations, while the outputs are quantified measures of a wing's BOS activity. Such

a relationship is reversed from what some may consider the normal concept of required activity level as the input and dollar obligations as the resultant cost, or output.

The analysis chapter which follows will sequentially revisit each of the three methodology steps in terms of their respective data gathered, analysis performed, and results achieved.

III. Analysis

Overview

This chapter discusses the data gathered, analysis performed, and results achieved from executing the methodology described in Chapter II. To review, this study's methodology calls for three main steps. First is the identification and selection of decision making units (DMUs) and their inputs and outputs via expert opinion, *a priori* judgement, and regression analysis. Second is the application of Data Envelopment Analysis (DEA) to identify sources and amounts of inefficiency, thus enabling formation of an efficient data set. And third is regression analysis of the efficient data, followed by comparing the results of regressing efficient versus inefficient data.

Identification and Selection of Inputs and Outputs

As mentioned in Chapter II, the author subjectively selected the Base Operations Support (BOS) function at each of 25 SAC bases as the DMUs for inclusion in this study. Names of the SAC bases are not used in this study, but rather the bases are labeled as DMU A through DMU Y.

Additionally, Chapter II refers to a comprehensive description of BOS components found at Appendix A. By way of review, the familiar activities which comprise BOS

are: base commander, comptroller, administration, supply, transportation, personnel, family services, chapel, judge advocate (legal), operations and training, safety, law enforcement, services, special services, and plans.

Identification and Selection of Inputs. To identify the most significant BOS components (from an O&M funds perspective), the author consulted with experts at HQ SAC/LGSMF, the primary headquarters funds manager for SAC BOS dollars. Their experience indicated that the areas of transportation, services, supply, and law enforcement were the major cost drivers within BOS obligations. Actual FY 85 BOS obligation data substantiated the experts' initial estimation, with transportation, services, supply, and law enforcement accounting for over 67% of the FY 85 3OS obligations managed by HQ SAC/LGSMF.

BOS Activity	FY 85 obligations	% of total
LGT (transportation)	$20.3 M	27.8
SV (services)	$17.6 M	24.0
LGS (supply)	$ 7.7 M	10.6
SP (law enforcement)	$ 3.8 M	5.2
		67.6

BOS funds managed by HQ SAC/LGSMF include the following four element of expense investment codes (EEICs) which represent the different types of items or services purchased: 473--rental of equipment, 569--purchased

maintenance of equipment, 592--miscellaneous contracts, and 6XX--supplies and equipment. Note that funding for civilian pay and military tempora y duty are not managed by HQ SAC/LGSMF.

Based on the above findings, the author selected FY 85 BOS obligations under EEICs 473, 569, 592, and 6XX for use as inputs for each of the 25 DMUs. This input variable is labeled OBLIG in this study. The required obligation data, as of 30 September 1985, was provided by HQ SAC/LGSMY.

Identification and Selection of Outputs: Subjective Analysis. With the selection of inputs complete, the next item involves identification of BOS outputs. The author employed three means in identifying BOS outputs: researching Air Force Manpower Standards (AFMSs) for workload factors, consulting functional experts, and using personal judgement.

Of the numerous workload factors listed in the AFMSs for the areas of transportation, services, supply, and law enforcement, the author deemed the following six measures (one is repeated) as straightforward and lending themselves well to this feasibility study:

Transportation

1. Monthly average number of registered vehicles assigned to the base.

2. Total base population--The sum of assigned officer, enlisted, and civilian government employees on a base.

Services

3. Military population--The officer and enlisted personnel assigned to a base.

4. Enlisted bed spaces--The number of housing assets available for permanent party unaccompanied enlisted personnel assigned to a base.

Supply

5. Total monthly transactions--The sum of supply and equipment transactions through the base supply system.

Law Enforcement

6. Post equivalents--The total number of personnel required to man all law enforcement posts and duties on a base.

7. Total base population (see 2 above).

In addition to researching AFMSs, the author consulted functional experts to determine what other output measures were available. HQ SAC/LGTX offerred transportation factors such as percentiles for VDM (vehicles down for maintenance), VDP (vehicles down for parts), and VOC (vehicles out of commission); however, the author decided that such percentiles would more logically vary with factors such as terrain, climate, usage, and vehicle age, and therefore were not included for consideration in this study. HQ SAC/LGSMF recommended using the total number of supply and equipment line items authorized through the base supply systems as an additional supply related measure. There were no other additional measures recommended.

Finally, personal judgement by the author produced

three additional output measures for consideration:

Transportation

1. Miles driven by the vehicle fleet. This would allow comparison of actual vehicle use and resultant maintenance costs (such as oil changes) across all DMUs.

2. Vehicle equivalents (for use instead of number of registered vehicles). This measure would capture the number of vehicles and their cumulative maintenance complexity, instead of just the number of vehicles.

Services

3. Enlisted population. This would be an additional facet to the study's population variables, and would possibly be a good predictor of BOS services obligations since expenses associated with supporting the enlisted population such as food service, dormitory support, and laundry and dry cleaning are a services responsibility.

The following is a summary of the nine BOS output measures identified for possible use in this study. The primary criterion used in developing this initial list of outputs was that a change in output should logically result in a direct change in the required input level (reference Chapter II's discussion of guidelines for selecting appropriate inputs and outputs). After each output is the variable name assigned to that output.

Supply

1. Average monthly number of supply and equipment line items authorized (LINITMS).

2. Average monthly number of total supply and equipment transactions (TRANSX).

Transportation

3. Vehicle equivalents assigned to a base (VEHEQ).

4. Average monthly number of miles driven by a base's fleet of vehicles (FLTMI).

Law Enforcement

5. Post equivalents on a base (POSTEQ).

Services

6. Military population assigned to a base (MILPOP).

7. Enlisted population of a base (ENLPOP).

8. Enlisted bed spaces (ENLBEDS).

General

9. Total population (TOTPOP).

In the discussion that follows, each of the nine outputs mentioned above is discussed in terms of the data desired for this study, the data actually gathered, and the source of the output data.

LINITMS. Desired data was the number of supply and equipment line items authorized for each DMU for each month of FY 85. The data was provided by HQ SAC/LGSMY. Values were missing for DMU M across all months, and were missing for February and March 1985 across all DMUs. The average monthly number of line items authorized was computed for each DMU by summing the values reported and dividing that sum by the number of months when values were reported. The number of months of data available for each DMU ranged from five to ten months.

TRANSX. Desired data was the total number of base supply transactions for each DMU for each month of FY 85. The data was provided by HQ SAC/LGSMY. Values were missing for DMU M across all months, and were missing for January, February, and March 1985 across all DMUs. The average monthly number of transactions was computed for each DMU in the same fashion as for LINITMS above. The number of months of data available for each DMU ranged from four to eight months.

VEHEQ. Desired data was the total number of vehicle equivalents for each DMU as of 30 September 1985. The data, provided by HQ SAC/LGTV, was complete.

FLTMI. Desired data was the number of vehicle fleet miles for each DMU for each month of FY 85. Data provided by HQ SAC/LGTX was complete, though it was only available for calendar year 1985. The author assumed that substituting the first fiscal quarter of FY 86 for the first fiscal quarter of FY 85 would not substantially alter the behavior of FLTMI. A simple 12-month average was computed for each DMU.

POSTEQ. Desired data was the total number of post equivalents for each DMU as of 30 September 1985. The only such data available from HQ SAC/SPPP was relatively current data as of 14 May 1986. On the assumption that the number of law enforcement posts and duties is not subject to frequent change, the data provided was used.

MILPOP. Desired data was the total number of officers and enlisted assigned to each DMU as of 30 September 1985. Data provided by HQ SAC/ACCC (cost analysis) was complete.

ENLPOP. Desired data was the total number of enlisted assigned to each DMU as of 30 September 1985. Data provided by HQ SAC/ACCC was complete.

ENLBEDS. Desired data was the total number of enlisted bed spaces for each DMU as of 30 September 1985. Data provided by HQ SAC/ACCC was complete.

TOTPOP. Desired data was the sum of assigned officer, enlisted, and civilian government employees at each DMU as of 30 September 1985. Data provided by HQ SAC/ACCC was complete.

Table I summarizes the numerical values for the input and output variables described above showing the mean, median, and range of observations for each variable. In addition, Appendix C shows the complete data set for all DMUs. Note that complete data were not available for DMU M, therefore, it was deleted from the observation set. Thus, the number of observations was reduced to 24.

Identification and Selection of Outputs: Regression Analysis. Having used expert opinion and *a priori* judgement to arrive at the nine outputs discussed above, regression/correlation analysis was then individually

TABLE I

Summary of Input and Output Data

	Range	Mean	Median
<u>Input</u>:			
OBLIGS ($000)	4,254.9 - 1,819.9	2,930.4	2,923.2
<u>Outputs</u>:			
LINITMS	62,373 - 1,919	14,855	8,014
TRANSX	100,240 - 38,041	70,681	73,110
VEHEQ	2,247.3 - 913.7	1,519.0	1,364.2
FLTMI	669,772 - 182,609	359,249	302,060
POSTEQ	151 - 55	76.7	66
ENLPOP	8,645 - 2,237	3,713	3,457
ENLBEDS	2,440 - 786	1,546	1,572
MILPOP	11,948 - 2,583	4,456	3,974
TOTPOP	14,355 - 3,397	5,488	5,088

applied to each output to evaluate the assumption that each output (= independent variable) varied directly or correlated positively with the input OBLIG (= dependent variable). As shown below, the coefficient of determination or R^2 (the amount of dependent variable variation explained by the independent variable) was positive for all nine selected outputs and this indicated a positive, direct relationship between each output and the input.

Output	R^2
LINITMS	.0499
TRANSX	.0960
VEHEQ	.4626
FLTMI	.1743
POSTEQ	.1873
ENLPOP	.1014
ENLBEDS	.0206
MILPOP	.0716
TOTPOP	.0771

At this point, the process of building a regression model to explain BOS obligations, the dependent variable, in terms of BOS outputs, the independent variables, was begun by applying stepwise regression to the data. Initially, an F test significance level of 0.1 was selected as the criterion for variables entering the model (meaning there must be 90% confidence that the coefficients of

variables entered into the model are not equal to zero). As a result, VEHEQ was the only variable that entered, having a significance level of .0004. Next, the variable entry criterion was raised to a 0.2 significance level, yielding the following four-variable model:

Model's R^2 = .6167

<u>Variables</u> (in order of entry)

VEHEQ

POSTEQ

LINITMS

TRANSX

For the purpose of observing the significance level at which the remaining output variables would enter the model, successive stepwise regressions were run while progressively decreasing the significance level, first to 0.3, then to 0.4, and finally to 0.5. No new output variables entered the model during this sequence, thus leaving the four-variable model described above unchanged.

In an effort to confirm the present four-variable model, two additional regression analyses were performed, one using backward elimination and one using a maximum R^2 criterion. Both procedures arrived at the same four-variable model described above.

Having arrived at an initial model, the author pursued explanation of why the intuitively significant

population variables (i.e., ENLPOP, MILPOP, and TOTPOP) had been excluded from entry into the model. By reviewing the results of the backward elimination regression mentioned earlier, it was noted that MILPOP was the last variable to be removed prior to arrival at the present four-variable model. Similarly, review of the maximum R^2 regression revealed that the best five-variable model included the four variables in the present model plus MILPOP. Furthermore, close examination of the data set revealed that DMU T, as shown below, had exceptionally high values and was an outlier for all three population variables.

Variable	DMU T	Mean
ENLPOP	8,645	3,713
MILPOP	11,948	4,456
TOTPOP	14,355	5,488

Concluding that DMU T was not a member of the set of normal SAC base populations, the author decided to eliminate DMU T from the observations in order to achieve homogeneity in the intuitively important population variables, and to see if elimination of DMU T's extreme population values would allow MILPOP to enter the model as a significant variable.

After removing DMU T from the set of observations (now consisting of 23 DMUs), another stepwise regression

(significance level = 0.2) yielded the following results:

Model's R^2 = .6304

Variables (in order of entry)

VEHEQ

MILPOP

TRANSX

POSTEQ

As anticipated, removing DMU T from the data set allowed MILPOP to enter the model as a significant variable in explaining variation in BOS obligations. The variable displaced by MILPOP's entry, LINITMS, decreased in significance, and did not enter into a model until the maximum R^2 regression included it in its best six-variable model, whereupon LINITMS entered with significance of 0.2710.

Regression Model I. With the removal of DMU T, the resulting four variables in the model--VEHEQ, MILPOP, TRANSX, and POSTEQ--offered the intuitive appeal of respectively representing the previously identified major BOS areas of transportation, services, supply, and law enforcement. The following statistical data and discussion describe this model, hereafter referred to as Model I.

Model I

$$OBLIG = 3.0507 + .0077\ TRANSX + .5797\ VEHEQ + 5.4835\ POSTEQ + .2544\ MILPOP$$

F value	7.675
Prob > F	.0009
R^2 value	.6304

Variable	t	Prob > /t/	Tolerance
TRANSX	1.564	.1352	.8881
VEHEQ	2.193	.0417	.6828
POSTEQ	1.369	.1880	.8325
MILPOP	2.188	.0421	.7321

The above F statistic, which in multivariate regression tests the hypothesis that the coefficients of all the independent variables equal zero, rejects that hypothesis at the .0009 significance level (i.e., there is 99.91% confidence that the coefficients for all the independent variables do not equal zero). Additionally, the R^2 value indicates that the four variables in the model explain just over 63% of variation in the dependent variable, BOS obligations.

The above t statistics, which test the hypothesis that the coefficient of each independent variable taken individually is equal to zero, reject that hypothesis at the .8648 confidence level (i.e., one minus Prob > /t/) for TRANSX; at the .9581 confidence level for VEHEQ; at the .8120 confidence level for POSTEQ; and at the .9579

confidence level for MILPOP.

The tolerance values listed above measure collinearity among the four independent variables. The computation of tolerance values, which can range from one (indicating independence among variables) to zero (indicating high collinearity), is achieved by first determining the R^2 between each variable and all remaining variables taken together and then subtracting those R^2 values from one. For example, TRANSX has an R^2 of only .1119 with the remaining three variables taken together, so its tolerance value is .8881, or one minus .1119. The tolerance values for all four variables in Model I are relatively high, thus indicating the absence of any serious collinearity problem.

The data set used to develop Model I was next examined for influential outliers via evaluating studentized residuals and studentized deleted residuals (26:405,6) both of which indicate outliers with respect to the input axis; leverage values (26:402) which indicate outliers with respect to the output axis; and Cook's distance measures (25:407) which indicate outliers with respect to both axes combined.

Studentized residuals measure an observation's distance from the regression surface in terms of a ratio of the observation's raw residual to the residual's sampling error. The computed criterion for identifying

outliers was 1.330, the t statistic for a 90% confidence interval and 18 degrees of freedom. This criterion was met or exceeded by the absolute values for the following observations:

DMU	Studentized Residual
A	-1.7241
B	-1.3369
H	1.4712
O	1.8108
P	-1.5299
W	2.2442

Studentized deleted residuals measure an observation's distance from the regression surface when that observation is excluded from the regression analysis. The computed criterion for identifying outliers was 1.740, for a 90% confidence interval. This criterion was met or exceeded by the absolute values for the following observations:

DMU	Studentized Deleted Residual
A	-1.8338
D	1.9460
W	2.5700

Leverage values measure how far displaced a DMU's observed output (= independent variable) is from the

center of the other DMUs' observed outputs. The computed criterion for identifying outliers was .4348, and was met or exceeded by the following observations:

DMU	Leverage
A	.5128
W	.4636

Finally, for Cook's distance measures, the computed criterion for identifying influential outliers with respect to both the input and output axes (i.e., combined impact) was 0.9044, and was not met or exceeded by any observations. The following two observations, however, had Cook's distance measures relatively close to the criterion:

DMU	Cook's distance measure
A	.626
W	.871

Since the next highest Cook's distance measure was .155 for DMU H, it was evident that DMUs A and W, though not highly influential, were at least moderately influential outliers. This classification of DMUs A and W as the most influential outliers in the observation set was confirmed by noting their consistent appearance as outliers under all three of the prior outlier analyses (i.e., studentized residuals, studentized deleted residuals, and leverage).

Upon concluding that DMUs A and W were relatively influential outliers, the data set was examined to determine the cause of the outlying observations. It was discovered that DMUs A and W had exceptionally high POSTEQ values of 151 and 144 respectively, compared to the next highest value of 97.

At this point, the author decided to investigate the effects of eliminating the influence of these two outliers on the model. The two approaches selected for dealing with the outliers were: (1) to eliminate the POSTEQ variable from the model and (2) to retain the POSTEQ output variable but remove the two outlying DMUs (A and W) from the data set. The results of both approaches are described in the following paragraphs.

Regression Model II. Under the first approach to reducing the influence of the outliers, a regression analysis was performed using three of the four output variables in Model I (TRANSX, VEHEQ, and MILPOP) and eliminating the fourth output variable, POSTEQ. The following statistical data and discussion compare this model, hereafter referred to as Model II, with Model I.

Model II

$$\text{OBLIG} = 120.0731 + .0093\ \text{TRANSX} + .6726\ \text{VEHEQ} + .2636\ \text{MILPOP}$$

F value	9.187
Prob > F	.0006
R^2 value	.5919

Variable	t	Prob > /t/	Tolerance
TRANSX	1.900	.0727	.9410
VEHEQ	2.574	.0186	.7309
MILPOP	2.220	.0387	.7345

Both the F value and Prob > F value for Model II are improved relative to Model I's values, which were 7.675 and .0009, respectively. Additionally, Model II's R^2 value is slightly less than that of Model I's .6304, as would be expected after removing a variable that explained some variation in the dependent variable.

Based on the below comparison of t statistics between Models I and II, all three output variables of Model II show increased significance levels.

	Prob > /t/	
Variable	Model I	Model II
TRANSX	.1352	.0727
VEHEQ	.0417	.0186
MILPOP	.0421	.0387

Furthermore, the below comparison of tolerance values between Models I and II indicates that the removal os POSTEQ resulted in decreased collinearity (increased independence) among the remaining three output variables

of Model II.

	Tolerance	
Variable	Model I	Model II
TRANSX	.8881	.9410
VEHEQ	.6828	.7309
MILPOP	.7321	.7345

The data set used to develop Model II was then analyzed to investigate prior outliers, DMUs A and W, as well as any new outliers. Appendix D shows a comparison of the full sets of residual analysis statistics for Models I and II.

DMUs whose studentized residuals under Model II met or exceeded the absolute value criterion of 1.328 (90% confidence interval and 19 degrees of freedom) are listed below:

DMU	Studentized Residual
L	-1.4885
O	1.5118
P	-1.5926
R	-1.3694
W	2.5549

The only DMU whose studentized deleted residual under Model II met or exceeded the computed outlier criterion of 1.734 (90% confidence interval) is listed

below:

DMU	Studentized Deleted Residual
W	3.0693

Under Model II, there were no DMUs whose leverage values met or exceeded the computed outlier criterion of .3478.

Finally, Model II offered no Cook's distance measures which met or exceed the computed influential outlier criterion of .870. DMU W, however, did have the highest Cook's distance measure of .276, which was notably greater than the next highest measure of .160.

Summarizing the outlier analysis of Model II, DMU A, which was an outlier under Model I, appeared nowhere as an outlier under Model II. DMU W, the other outlier under Model I, was still an influential outlier relative to the input axis (reference its high studentized deleted residual under Model II). However, with leverage of only .1449 compared to the outlier criterion of .3478, DMU W was no longer an outlier relative to the output axis. Consequently, DMU W's combined outlier influence (Cook's distance measure) was significantly reduced from .871 under Model I to .276 under Model II. Additionally, no new influential outliers were introduced using Model II.

<u>Regression Model III</u>. The second approach to reducing the influence of outliers was to eliminate DMUs

A and W from the observation set while retaining all four output variables of Model I. The following statistical data and discussion compare results of this approach, hereafter referred to as Model III, with Model I.

Model III

$$OBLIG = 310.695 + .0116\ TRANSX + .3359\ VEHEQ - 7.0927\ POSTEQ + .4042\ MILPOP$$

F value	7.939
Prob > F	.0010
R^2 value	.6650

Variable	t	Prob >/t/	Tolerance
TRANSX	2.432	.0272	.8198
VEHEQ	1.038	.3146	.4384
POSTEQ	-0.620	.5439	.4329
MILPOP	2.486	.0243	.2993

Comparing Model III's F and Prob > F values with those of Model I (7.675 and .0009 respectively) reveals only slight change. Similarly, Model III's R^2 value is only slightly greater than Model I's R^2 of .6304. Significant difference between the two models is evident, however, from comparison of the remaining statistics.

Comparison of t statistics between Models I and III yields mixed results. As shown below, the significance of variables TRANSX and MILPOP increases, while the significance of variables VEHEQ and POSTEQ decreases

under Model III.

	Prob >/t/	
Variable	Model I	Model III
TRANSX	.1352	.0272
VEHEQ	.0417	.3146
POSTEQ	.1880	.5439
MILPOP	.0421	.0243

In fact, significance levels of VEHEQ and POSTEQ no longer meet the 0.2 significance level criterion originally used to select variables for Model I. Furthermore, the removal of DMUs A and W from the observation set (along with their extreme POSTEQ values) rendered POSTEQ an insignificant explanatory variable for variation in BOS obligations. Note the confidence level that POSTEQ's coefficient is not in fact zero is a low 45.6% (i.e., 1 - .5439 = .4561).

Comparison of tolerance values also indicates notable difference between Models I and III. As shown below, the tolerance values of all four variables decrease with Model III, indicating an increase in collinearity.

	Tolerance	
Variable	Model I	Model III
TRANSX	.8881	.8198
VEHEQ	.6828	.4384
POSTEQ	.8325	.4329
MILPOP	.7321	.2993

An additional characteristic of Model III worthy of critique, and which was present in neither Models I nor II, is that Model III assigns a negative coefficient to POSTEQ. The assignment of a negative coefficient to an output variable contradicts the earlier mentioned primary *a priori* criterion for identifying and selecting outputs--that they should have a direct (not inverse) relationship with input.

Since the weaknesses of Model III discussed above have already identified it as a nonviable model, outlier analysis under Model III was not pursued.

Summarizing results from the above two approaches for reducing the influence of outliers found in Model I, Model II with its improvement in all areas except for a slight reduction in R^2 was judged superior to Model III with its decreases in variable coefficient significance and increases in collinearity. The author therefore selected Model II, with output variables TRANSX, VEHEQ, and MILPOP, as this study's illustrative, yet reasonably valid model for predicting BOS input as a function of BOS outputs.

Later in this chapter, input predictions generated by Model II will be compared with efficient input predictions, generated by regression of an efficient data set. The obtainment of that efficient data set is the next subject of discussion and the second main step in

this study's methodology.

Application of Data Envelopment Analysis

In this step DEA was applied to the original input and output data (OBLIG, TRANSX, VEHEQ, and MILPOP) which was used to develop Model II, discussed above. The results of DEA, shown in Table II, include both the relative efficiency measures, referred to as h_0 in Appendix B, and output slack (= shortages), referred to as s_r^+ in Appendix B, for each DMU. Of the 23 DMUs involved, five were rated efficient (h_0 = 1.000), while the remaining 18 inefficient DMUs held efficiency measures ranging from .9726 to .6106. Using DMU L as an example, its rating of .9726 indicates that the DMU was 97.26% efficient and should be able to reduce obligations by 2.74%. In addition, the output shortage of 623 for MILPOP indicates that even with this 2.74% reduction in obligations, DMU L should be able to support 623 additional military personnel.

When considering the magnitude and meaning of the output shortages listed in Table II, it is important to bear in mind that they are not intended to be used for management control purposes, but rather for estimating purposes. For instance, referring again to DMU L, the output shortage of 623 under MILPOP does not mean that management should increase base military population by 623. Rather, it means that based on the production

Table II

Results of Data Envelopment Analysis (DEA)

DMU	h_0	Output Shortages (slack)		
		TRANSX	VEHEQ	MILPOP
A	.8974	-	-	1202
B*	1.0000	-	-	-
C	.8548	-	549	-
D	.9045	-	193	-
E*	1.0000	-	-	-
F	.9263	7594	381	-
G	.9606	-	310	-
H	.7427	-	-	-
I	.8503	-	150	-
J	.8131	1898	-	-
K	.8298	-	-	-
L	.9726	-	-	623
N	.7499	-	8	-
O	.6872	-	-	896
P*	1.0000	-	-	-
Q	.9595	-	464	-
R*	1.0000	-	-	-
S	.8864	20,712	-	-
U	.8036	-	120	-

* Efficient DMUs

Table II (continued)

DMU	h_0	Output Shortages (slack) TRANSX	VEHEQ	MILPOP
V	.8572	-	163	-
W	.6106	-	-	596
X	.7136	-	-	-
Y*	1.0000	-	-	-

* Efficient DMUs

possibilities curve derived from existing, efficient, similar DMUs, DEA estimates that DMU L could support an additional 623 military personnel without an increase in BOS funding input. In fact, it should be able to support these additional personnel with 2.74% fewer obligations.

Using the DEA results shown in Table II, it was possible to transform the original (inefficient) set of inputs and outputs into a set of efficient inputs and outputs. Recalling the computation of efficient values described at the end of Appendix B, efficient output levels are computed by adding output shortages to the observed amounts of output, while efficient input levels are computed by multiplying the efficiency measures (h_0) times the observed amounts of input.

Using DMU L as an example, the observed TRANSX and VEHEQ amounts are already efficient. For MILPOP,

however, the efficient amount is computed by adding the shortage of 623 to the observed amount of 4514, giving 5137. The efficient input level is computed by multiplying the observed input of 3003 by .9726, giving 2920.7.

Since the present study involves only one input, there are no input slack values to include in the computation of efficient inputs. At Appendix E is the set of efficient inputs and outputs computed using the results of DEA shown earlier in Table II.

With the computation of the efficient data set, the second step of this study's methodology was complete, making way for the third and final part of this analysis.

Regression Analysis of Efficient Data

The final step of this study's methodology was to perform regression analysis on the efficient inputs and outputs computed above, and then to compare the results of regressing efficient data with those of regressing inefficient data. The following discussion of this topic has two parts. First, the statistics of the efficient regression model, hereafter referred to as Model E, will be compared with those of the inefficient regression model--the previously described Model II. And second, the predicted input values generated by Model E (efficient) and Model II (inefficient) will be compared to detect the impact of using efficient inputs and outputs on predicting

BOS funding levels.

As shown below, the F value, Prob > F, and R^2 of Model E show impressive improvement over Model II.

Model E

$$OBLIG = 194.5352 + .0071\ TRANSX - .0529\ VEHEQ + .4333\ MILPOP$$

	Model II	Model E
F value	9.187	201.427
Prob > F	.0006	.0001
R^2 value	.5919	.9695

Model E's larger F value indicates that with the efficient data set, input values have much less variability around their regression surface than is true for the inefficient data. Bearing in mind that the F value is the ratio of the regression mean square (MSR) to the error mean square (MSE) (26:92), Model E's larger F value is explained by comparing its MSE of 6661.6 with Model II's MSE of 195,766.9. Model E's use of efficient data reduced the MSE to less than one twenty-ninth of its magnitude with inefficient data. Model E's increased R^2 further highlights the effect of this reduced variation among efficient data points.

From the below comparison of t statistics between Model II and Model E, it is evident that TRANSX and MILPOP increase, while VEHEQ decreases in significance

Model II

Variable	t	Prob > /t/
TRANSX	1.900	.0727
VEHEQ	2.574	.0186
MILPOP	2.220	.0387

Model E

Variable	t	Prob > /t/
TRANSX	6.736	.0001
VEHEQ	-0.623	.5405
MILPOP	14.596	.0001

In addition to VEHEQ losing significance in Model E, the coefficient for VEHEQ changes signs from positive (in Model II) to negative (in Model E). The explanation for these changes becomes evident in the following discussion of tolerance values.

The below comparison of tolerance values between Model II and Model E indicates that the use of efficient inputs and outputs in Model E introduced collinearity among the output variables.

Tolerance

Variable	Model II	Model E
TRANSX	.9450	.8338
VEHEQ	.7309	.3455
MILPOP	.7345	.3849

Though all three output variables suffered loss of independence, it was particularly true for VEHEQ and MILPOP. Note that the low tolerance of .3455 for VEHEQ means that 65.45% of variation in VEHEQ data is explained by variations in TRANSX and MILPOP. The previously mentioned loss of significance and change in sign for VEHEQ are typical symptoms of collinearity. A possible explanation for this occurrence is that operational inefficiencies reflected in Model II masked the collinearity problem which became apparent when these inefficiencies were removed in Model E.

Having completed comparison of the two models' statistics, the second part of this final analysis step is to compare the predicted input values generated by Model II and Model E to determine the impact of using efficient data on predicting BOS obligation levels.

Table III shows the following five columns of data for each DMU: "FY 85 Actual", the actual BOS obligations for FY 85; "Model II Predicted", obligations predicted by Model II using inefficient (raw) data; "Model E Predicted", obligations predicted by Model E using efficient data; "Model E minus Actual", the difference between predicted efficient obligations and actual obligations; and "Model E minus Model II", the difference between predicted efficient obligations and predicted inefficient obligations. Negative amounts in columns

Table III

Comparison of Predicted Inputs

($000)

DMU	(1) FY 85 Actual	(2) Model II Predicted	(3) Model E Predicted	(4) Model E minus Actual	(5) Model E minus Model II
A	3359.1	3564.2	3058.7	(300.4)	(505.5)
B	3032.0	3464.8	3098.6	66.6	(366.2)
C	3244.9	2857.2	2625.4	(619.5)	(231.8)
D	2127.8	2143.7	1879.5	(248.3)	(264.2)
E	2584.2	2728.1	2562.4	(21.8)	(165.7)
F	2923.2	2525.4	2675.6	(247.6)	150.2
G	2972.2	3105.6	2932.4	(39.8)	(173.2)
H	4254.9	3836.8	3102.1	(1152.9)	(734.7)
I	2865.4	3036.4	2537.5	(327.9)	(498.9)
J	2459.1	2477.1	1989.3	(469.8)	(487.8)
K	3595.7	3832.8	3033.6	(562.1)	(799.2)
L	3003.0	3600.9	2918.5	(84.5)	(682.4)
N	3105.8	2788.1	2282.6	(823.2)	(505.5)
O	3565.8	2940.4	2555.4	(1010.4)	(385.0)
P	2050.4	2680.3	2021.9	(28.5)	(658.4)
Q	2471.7	2683.4	2430.6	(41.4)	(252.8)

( ) Negative amount means a reduction is needed to equal the predicted efficient input.

Table III (continued)

($000)

DMU	(1) FY 85 Actual	(2) Model II Predicted	(3) Model E Predicted	(4) Model E minus Actual	(5) Model E minus Model II
R	1819.9	2390.2	1956.3	136.4	(433.9)
S	3581.5	3463.3	3068.7	(512.8)	(394.6)
U	2824.8	2667.3	2245.4	(579.4)	(421.9)
V	2935.3	2921.8	2466.9	(468.4)	(454.9)
W	4237.7	3192.3	2565.3	(1672.4)	(627.0)
X	2429.6	2234.0	1812.9	(616.7)	(421.1)
Y	2167.4	2477.1	2043.2	(124.2)	(433.9)
	67611.4			(9748.6)	(9748.4)

( ) Negative amount means a reduction is needed to equal the predicted efficient input.

four and five indicate that reduction of amounts in columns one and two would be necessary to arrive at the predicted efficient obligations. The opposite is true for positive amounts in columns four and five. As indicated at the bottom of Table III, the sum of efficient obligations for all DMUs represents a $9.7 million (14.4%) reduction from the $67.6 million sum of actual FY 85 obligations. Stated differently, if all DMUs had been operating efficiently, FY 85 BOS obligations would

have been $9.7 million less. Additionally, using predicted efficient inputs rather than actual FY 85 obligations for each DMU as a funding baseline for FY 86 would reduce that baseline by the same $9.7 million.

Analysis Summary

In summary, this chapter has discussed the data gathered, analyses performed, and results achieved during the process of identifying and selecting i puts and outputs, applying Data Envelopment Analysis, and performing regression analysis of the resultant efficient data. Key results of the preceding analysis include the following three items. First, Model II, with output variables TRANSX, VEHEQ, and MILPOP, was selected as an illustrative yet reasonably valid model for predicting BOS obligations as a function of BOS output. Second, a data set of efficient inputs and outputs was computed using efficiency measures and output shortages identified by DEA. And finally, regression analysis of that efficient data generated a set of predicted efficient input levels for all DMUs, the sum of which was $9.7 million less than actual (inefficient) FY 85 BOS obligations.

The conclusions and recommendations chapter which follows will distill the significant findings of this research, and make recommendations for areas of further study.

IV. Conclusions and Recommendations

Introduction

As introduced in Chapter I, the reason for this research is to respond to the call for efficient use of federal funding by studying the feasibility of a technique designed to facilitate efficient allocation of O&M funds. More specifically, the objective of this research is to determine the feasibility of using Data Envelopment Analysis (DEA) in combination with regression analysis to provide a method for indicating efficient allocation of Air Force O&M funds. The proposed approach uses expert opinion, *a priori* judgement, and regression analysis to identify outputs which have a significant impact on BOS obligations. Then, using these outputs and BOS obligations as the input, DEA is initiated to determine efficient output and obligation levels. Finally, regression analysis is performed on the data set of efficient outputs and obligations in order to predict an efficient BOS obligations or funding level for each SAC wing involved.

Conclusions

Model II. Model II, with output variables TRANSX, VEHEQ, and MILPOP, was selected as this study's illustrative, yet reasonably valid model for predicting BOS input

as a function of BOS outputs. The model (which is illustrative--not exhaustive) is based on data limited to one MAJCOM, one program element, and four elements of expense, and therefore attempts to capture only a small, representative sample of the vast Air Force O&M program. The author's satisfaction with Model II's modest R^2 of .5919 further emphasizes that the purpose of this study was not to derive a rigorous model for explaining BOS obligations, but rather an illustrative model for use as a reference when analyzing the effect of subsequent regression of efficient data.

Additionally, differences between the originally desired and actually gathered output data, as described early in Chapter III, may have biased the ability of certain outputs to explain variation in BOS obligations. For instance, of the twelve months in FY 85, availability of data for each DMU varied from four to eight months for TRANSX, and from five to ten months for LINITMS.

Model E. Model E is the product of performing regression analysis on the efficient input and output values computed using the efficiency measures and output shortages indicated by DEA. Model E has significantly less variability of data about its regression surface than does Model II. Accompanying the reduced variability of efficient input and output data, however, is the introduction of collinearity among outputs (as evidenced by

low tolerance, sign change, etc.). As Chapter III mentions, this could possibly be the result of removing the effect of operational inefficiencies which had previously masked this problem in the data. Since the expert opinions and a priori judgement used in output selection naturally assumes a causal link between input and outputs; and since Model II verifies the ability of TRANSX, VEHEQ, and MILPOP to explain BOS obligations, the final model despite its collinearity should still be useful as long as the causal (or even statistical) relationship between independent variables is expected to remain constant. Stated more succinctly,

> [T]he presence of serious multicollinearity often does not affect the usefulness of the fitted model for making inferences about mean responses or making predictions, provided that the values of the independent variables for which inferences are to be made follow the same multicollinearity pattern as the data on which the regression model is based [26:393].

Concept Application. Briefly recapping Chapter I's opening discussion of this study's problem statement: the current method for allocating O&M funds involves a distribution based on an entity's historical spending, plus or minus amounts corresponding to changes in that entity's mission. The problem with the current method is that historical spending inefficiencies may get carried into the baseline for future funding allocation.

The concept of Model E's predicted efficient input values offers a potential alternative to using inefficient

prior fiscal year obligations as the baseline for distribution of current year funds. Assuming functional experts and budgeteers could agree on a set of appropriate output measures (i.e., ones that are representative of organizational goals and objectives, logically and directly related to funding inputs, not susceptible to alteration or unethical manipulations, and auditable), a process of efficiency based budgeting could evolve.

Somewhat analogous to the concept of Zero Based Budgeting, funds distribution under efficiency based budgeting would require justification of funding increments above a DMU's predicted efficient input (as opposed to the zero reference of ZBB). Though specifics on implementing such a process are beyond the scope of this paper, the potential for reducing inefficient spending via efficient allocation of funds is evident by the $9.7 million of theoretically inefficient spending identified by this study.

Recommendations

As a result of this research, the author recommends four areas requiring further study. First, before efficient allocation of funds can move from concept to practice, appropriate output measures need to be designed that will enable accurate analysis of efficient input requirements. Such a framework of output measures, which would link funding cuts to decreased output and funding

increases to increased output, would enhance the Air Force's financial management control system and strengthen the credibility of funds justifications.

Second, in specifying a regression model depicting BOS obligations as a function of BOS outputs, the use of indicator variables to differentiate between bomber and missile wings may enhance the model's ability to explain variation among observed DMUs.

Third, to better analyze and compare relative efficiency among DMUs, the efficiency measures (h_0) generated by this study's DEA might be broken down into components of technical and scale efficiency. Such a distinction would, for instance, allow an analyst to determine that an inefficient DMU ($h_0 < 1.0$) was in fact operating with technical efficiency, but had a physical plant which was scale inefficient.

Finally, the fourth area recommended for further study is the problem of introducing collinearity into a regression model when using efficient data. As mentioned above in the concluding remarks about Model E, serious collinearity does not necessarily render a model useless. It would be informative, however, to investigate a remedial measure such as ridge regression (NWK:394) as a way of dealing with the collinearity introduced by efficient output data.

Appendix A:

Description of SAC Base Operations Support Activity

Below is the description of BOS activity for strategic offensive forces. The description of BOS activity for strategic defensive forces (21:M-00,01) is not listed here since it is nearly identical to that for offensive forces.

Base Operations - Offensive: Includes manpower authorizations, peculiar and support equipment, necessary facilities and the associated costs specifically identified and measurable to the following functional categories of effort at installations having a primary mission of supporting strategic offensive forces: (1) Administration - Installation Headquarters Administration and Command (Including squadron level responsible for Base Operations), Installation Comptroller Services, Installation ADP Services, Installation Information Activities, Installation Legal Activities, Installation Civilian Personnel Administration, Installation Military Personnel Administration, Installation Safety, Installation Management Analysis/Management Engineering, Retail Supply Operations, Installation Storage Activities; (2) Maintenance of Installation Equipment (Includes maintenance of administrative aircraft, vehicles and equipment but excludes maintenance of tactical equipment, combat vehicles and mission aircraft); (3) Other Base Services - Installation Transportation Activities, Installation Training (Excludes Troop Training and Tactical Exercises), Installation Physical Security and Police Activities, Laundry and Dry Cleaning (For Troop Support and other Appropriated Fund Activities), Installation Airfield/Air Base Operations (Control Tower, Weather, Flight Services, etc.), Installation Restoration Bachelor Housing Operations and Furnishings (Management; Housing Assignment; Care of Quarters; Provisions, Care, Preservation and Maintenance of Furnishings, etc.); (4) Other

Personnel Support - Food Service, Social Actions, Community Services, Chaplains, Bands, Morale, Welfare and Recreation. Excludes the following functional categories which are operating support, but are reported under seperate PEs: Maintenance and Repair of Real Property, Minor Construction, Operation of Utilities, Other Engineering Support, Base Communications, Commissary Operations: Retail and Troop Issue Stations, Hospitals, Medical and Dental Clinics and Dispensaries, Family Housing [21:J-17,18].

Appendix B: DEA Model

DEA in its fractional programming format is as follows:

$$\text{Maximize:} \qquad h_0 = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}} \tag{1}$$

$$\text{Subject to:} \qquad 1 \geq \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \tag{2}$$

$$j = 1, \ldots n$$

$$u_r, v_i \geq \epsilon > 0$$

where

h_0 = the efficiency measure for DMU "0" which is the DMU being measured relative to the other DMUs.

x_{i0} = the observed amount of the ith input used by DMU "0" during the observed period.

y_{r0} = the observed amount of output "r" that DMU "0" produces during the observed period.

x_{ij} = the observed amount of the ith input that DMU "j" produces during the observed period.

v_i and u_r = values the model determines directly from the data to be used in the function.

ϵ = a small, positive non-Archimedian constant [24:42].

The DEA model's mathematical programming operates under two conditions. The constraint equation guarantees the first condition of the model which states that:

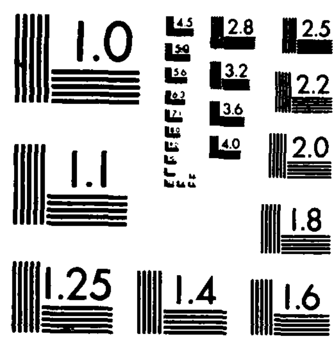
> outputs and inputs are to be weighted in such a way that each unit is compared to all of the others in the set and is constrained not to be larger than the best input/output ratio observed for any DMU [5:59].

The objective function guarantees the second condition of the model which states that "weights are calculated to give the largest possible ratio value for a given unit without violating the first condition" (5:59).

The objective function is the ratio of a DMU's weighted multiple outputs to its weighted multiple inputs. The "y" and "x" values are not variables, but rather are known constants--actual measures of output and input, respectively. The variables "u" and "v" in the objective function are the output and input weights respectively, which DEA solves for by comparing each DMU's output-to-input ratio with those of all other DMUs in the data set (5:60). To avoid confusion with normal use of <u>a priori</u> weights Charnes, Cooper, and Rhodes (13) refer to these variables as transformation ratios. This terminology refers to the fact that they

AD-A174 371 FEASIBILITY OF USING DATA ENVELOPMENT ANALYSIS AND REGRESSION ANALYSIS IN (U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF SYST J R WALLACE SEP 86 AFIT/GSM/LSY/86S-21 2/2

UNCLASSIFIED F/G 5/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

"transform real inputs (x_{i0}) to a 'virtual' input (X_0) and real outputs (y_{r0}) to a 'virtual' output (Y_0)" (7:78). The resulting weights or transformation ratios of each DMU are used to determine the DMU's measure of relative efficiency.

The fractional programming model described above conceptualizes DEA's measurement of relative efficiency; however, the actual computerized mathematical operations with DEA take the form of an ordinary linear programming model, capable of indicating slack (excess resource) in both outputs and inputs (5:61). The linear programming format shown below has been proven by Charnes, Cooper, and Rhodes (13) to be an equivalent mathematical transformation of the previously shown fractional programming model.

Minimize: $$h_0 = \theta - \epsilon \left(\sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^- \right) \qquad (3)$$

Subject to: $$\sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+ = y_{r0} \qquad (4)$$

$r = 1, \ldots s;$

$$-\sum_{j=1}^{n} x_{ij} \lambda_j - s_i^- + \theta x_{i0} = 0 \qquad (5)$$

$i = 1, \ldots, m$

$\lambda_j,\ s_r^+,\ s_i^- \geq 0$

θ unrestricted in sign

where

θ = the intensity multiplier of input x_{i0}

λ = the variable determined by the model,

s_r^+ = output slack value for output "r",

s_i^- = input slack value for input "i",

ϵ = a small, positive non-Archimedian constant [24:43].

At this point, with a DMU's efficiency measure and slack values determined, its efficient levels of output and input are computed as follows (13):

$$\hat{y}_r = y_r + s_r^+ \tag{6}$$

$$\hat{x}_i = h_0 x_i - s_i^- \tag{7}$$

where

$\hat{y}_r$ = the efficient output level for output "r",

$\hat{x}_i$ = the efficient input level for input "i",

y_r, s_r^+, h_0, x_i, s_i^- are as previously defined.

Appendix C: Observed Data

	Input ($000)	Outputs			
DMU	OBLIG	LINITMS	TRANSX	VEHEQ	FLTMI
A	3359.1	11328	97747	2169.8	373871
B	3032.0	29369	73110	1719.7	302060
C	3244.9	20553	100240	1047.9	282722
D	2127.8	5327	69569	913.7	197353
E	2584.2	16139	43552	1347.3	290827
F	2923.2	62373	38041	1030.4	240641
G	2972.2	7059	72293	1345.0	222902
H	4254.9	32375	72714	2247.3	669772
I	2865.4	6298	86322	1495.2	313069
J	2459.1	2920	40586	1518.5	557474
K	3595.7	6767	96784	2129.6	639409
L	3003.0	10740	86311	2207.4	365648

Appendix C (Continued)

DMU	Outputs				
	POSTEQ	ENLPOP	ENLBEDS	MILPOP	TOTPOP
A	151	3668	1344	4065	5088
B	97	4753	2440	5710	7199
C	94	3614	1070	4157	4933
D	57	2479	786	2880	3397
E	81	4060	1206	4913	6153
F	70	4261	959	5148	5862
G	78	4480	2231	5332	6032
H	67	4880	1387	5789	6615
I	80	3605	1572	4189	5022
J	62	3003	1700	3629	4329
K	64	4451	1910	5221	6025
L	59	3730	1227	4514	7657

Appendix C (Continued)

	Input ($000)	Outputs			
DMU	OBLIG	LINITMS	TRANSX	VEHEQ	FLTMI
N	3105.8	9158	83775	1387.6	252930
O	3565.8	9279	84880	1695.8	237231
P	2050.4	3987	46634	1723.9	652451
Q	2471.7	43576	84401	1086.0	258924
R	1819.9	8014	64383	1242.8	278913
S	3581.5	47704	40986	2099.1	656076
T	2717.5	2241	43737	1343.6	308659
U	2824.8	5767	81528	1259.1	220868
V	2935.3	5890	90375	1364.2	318592
W	4237.7	1964	71454	2017.9	487844
X	2429.6	1919	48164	1260.3	485673
Y	2167.4	5787	78748	1189.0	184716

Appendix C (Continued)

	Outputs				
DMU	POSTEQ	ENLPOP	ENLBEDS	MILPOP	TOTPOP
N	66	3095	1304	3612	4338
O	63	2942	2358	3364	4244
P	63	3091	1664	3661	4354
Q	73	3234	1824	3962	5209
R	61	2663	1196	3159	3889
S	68	5011	2132	5875	6768
T	126	8645	1914	11948	14355
U	61	3132	1920	3561	4305
V	65	3457	1937	3945	4633
W	144	3002	822	3974	5810
X	56	2623	864	3097	3762
Y	57	2699	1214	3117	3724

Appendix D: Statistics for Analyzing Residuals of Models I and II

	Studentized Residual		Studentized Deleted Residual		Leverage		Cook's Distance Measure	
DMU	Model I	Model II	Model I	Model II	Model I	Model II	Model I	Model II
A	-1.7241	-0.5278	-1.8338	-0.5175	0.5128	0.2288	0.626	0.021
B	-1.3369	-1.0772	-1.3690	-1.0821	0.1963	0.1754	0.087	0.062
C	0.7843	1.0204	0.7756	1.0216	0.2923	0.2625	0.051	0.093
D	0.0288	-0.0397	0.0280	-0.0386	0.1774	0.1753	0.000	0.000
E	-0.5954	-0.3577	-0.5844	-0.3493	0.1949	0.1732	0.017	0.007
F	0.9524	1.0972	0.9498	1.1035	0.3399	0.3286	0.093	0.147
G	-0.3954	-0.3336	-0.3859	-0.3257	0.1853	0.1840	0.007	0.006
H	1.4712	1.0682	1.5243	1.0724	0.2641	0.2174	0.155	0.079
I	-0.4247	-0.4011	-0.4148	-0.3921	0.0717	0.0716	0.003	0.003
J	-0.0043	-0.0457	-0.0042	-0.0445	0.2074	0.2066	0.000	0.000
K	-0.1925	-0.6022	-0.1872	-0.5918	0.2873	0.2079	0.003	0.024
L	-1.1326	-1.4885	-1.1422	-1.5414	0.2662	0.1756	0.093	0.118

Appendix D (Continued)

	Studentized Residual		Studentized Deleted Residual		Leverage		Cook's Distance Measure	
DMU	Model I	Model II	Model I	Model II	Model I	Model II	Model I	Model II
N	0.8915	0.7473	0.8862	0.7383	0.0845	0.0769	0.015	0.012
O	1.8108	1.5118	1.9460	1.5689	0.1521	0.1259	0.118	0.082
P	-1.5299	-1.5926	-1.5941	-1.6653	0.2056	0.2010	0.121	0.160
Q	-0.5529	-0.5143	-0.5419	-0.5041	0.1349	0.1345	0.010	0.010
R	-1.3285	-1.3694	-1.3595	-1.4040	0.1168	0.1142	0.047	0.060
S	0.4937	0.3206	0.4831	0.3129	0.3155	0.3057	0.022	0.011
U	0.5354	0.3720	0.5246	0.3634	0.0954	0.0842	0.006	0.003
V	0.2005	0.0320	0.1951	0.312	0.1112	0.0979	0.001	0.000
W	2.2442	2.5549	2.5700	3.0693	0.4636	0.1449	0.871	0.276
X	0.5880	0.4920	0.5770	0.4819	0.1959	0.1929	0.017	0.014
Y	-0.5944	-0.7441	-0.5834	-0.7351	0.1288	0.1149	0.010	0.018

Appendix E: Efficient Data

DMU	Input ($000) OBLIG	Output TRANSX	VEHEQ	MILPOP
A	3014.5	97747	2169.8	5268
B	3032.0	73110	1719.7	5710
C	2773.7	100240	1597.4	4157
D	1924.6	69569	1106.4	2880
E	2584.2	43552	1347.3	4913
F	2707.8	45635	1411.7	5148
G	2855.1	72293	1654.8	5332
H	3160.1	72714	2247.3	5789
I	2436.4	86322	1645.7	4189
J	1999.5	42484	1518.5	3629
K	2983.7	96784	2129.6	5221
L	2920.7	86311	2207.4	5137
N	2329.0	83775	1396.1	3612
O	2450.4	84880	1695.8	4260
P	2050.4	46643	1723.9	3661
Q	2371.6	84401	1549.8	3962
R	1819.9	64383	1242.8	3159
S	3174.6	61698	2099.1	5875
U	2270.0	81528	1379.0	3561
V	2516.1	90375	1527.6	3945
W	2587.5	71454	2017.9	4543
X	1733.8	48164	1260.3	3097
Y	2167.4	78748	1189.0	3117

Bibliography

1. Alderman, Karen. "Efficiency Reviews and Operational Improvements--the Key to Future Effectiveness," Defense Management Journal, 9: 12-17 (Third Quarter 1983).

2. Banker, R. and others. "A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations From a Known Technology." Working Paper 85-08-2, Carnegie-Mellon University, January 1984 (revised March 1985).

3. Banker, R. and Others. "Some Models for Estimating Technical and Scale Inefficiency in Data Envelopment Analysis," Management Science, 30: 1078-1092 (September 1984).

4. Banker, R. "Bi-extremal Principle for Frontier Estimation and Efficiency Evaluations," Management Science, 27: 1370-1382 (December 1981).

5. Bessent, Authella M. and E. Bessent. "Determining the Comparative Efficiency of Schools Through Data Envelopment Analysis," Educational Administration Quarterly, 16: 57-75 (Spring 1980).

6. Bowlin, William F. and others. "Data Envelopment Analysis and Regression Approaches to Efficiency Estimation and Evaluation," Annals of Operations Research, 2: 113-138 (1985).

7. Bowlin, William F. A Data Envelopment Analysis Approach to Performance Evaluation in Not-For-Profit Entities With an Illustrative Application to the U.S. Air Force. PhD Dissertation. University of Texas, Austin TX, December 1984.

8. Byers, Timothy and Maj Susanne M. Waylett. Measuring of Air Force Fire Department Productivity: An Evaluation of Efficiency/Effectiveness. MS Thesis GEM/LSM/84S-4. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1984 (AD-A146874).

9. Byrne, Vincent J. "Futurism in Resource Management: the Quest for Productivity Through Organizational Effectiveness," Armed Forces Comptroller, 23: 7-9 (November 1978).

10. Charnes, A. and others. "Pareto-optimality, Efficiency Analysis, and Empirical Production Functions," Journal of Econometrics, (submitted).

11. Charnes, A. and W. Cooper. "Preface to Topics in Data Envelopment Analysis," Annals of Operations Research, 2: 59-94 (1985).

12. Charnes, A. and others. "A Developmental Study of Data Envelopment Analysis in Measuring the Efficiency of Maintenance Units in the U.S. Air Forces," Annals of Operations Research, 2: 95-112 (1985).

13. Charnes, A. and others. "Measuring the Efficiency of Decision Making Units," European Journal of Operational Research, 2: 429-444 (November 1978).

14. Clark, C.T. Lecture materials distributed in LOGM 635, Evaluating the Efficiency, Productivity and Performance of Air Force Organizations. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, January 1986.

15. Comptroller General of the United States. The Defense Budget: A Look at Budgetary Resources, Accomplishments, and Problems. Report GAO/PLRD-83-62. Washington: U.S. Government Printing Office, 27 April 1983.

16. Congressional Budget Office and General Accounting Office. Analysis of the Grace Commission's Major Proposals for Cost Control: A Joint Study by the Congressional Budget Office and General Accounting Office. Washington: U.S. Government Printing Office, February 1984.

17. Cooper, W. and Y. Ijiri. Kohler's Dictionary For Accountants (Sixth Edition). Cliffs NJ: Princeton Hall, Inc., 1983.

18. Department of the Air Force. Comptroller Activities, Functions, and Responsibilities. AFR 170-6. Washington: HQ USAF, 6 August 1982.

19. Department of the Air Force. USAF Budget Manual: Budget Management for Operations. AFM 172-1, Volume III. Washington: HQ USAF, 30 September 1983.

20. Department of the Air Force. The Air Force Budget Process. AFP 172-4. Washington: HQ USAF, 1 October 1984.

21. Department of the Air Force. Air Force Data Dictionary. AFR 700-20. Volume I, Part V. Washington: HQ USAF, 1 August 1985.

22. Hitt, Capt Richard E. and Maj Robert F. Horace. Feasibility of Measuring Technical Productivity Improvements in Air Force Logistics Command Depot-Level Maintenance Using the Data Envelopment Analysis (DEA) and Constrained Facet Analysis (CFA) Models. MS Thesis GLM/LSM/84S-30. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1984 (AD-A147329).

23. Lein, P.S. "30 Years of Management Analysis," The Air Force Comptroller, 13: 28-31 (July 1979).

24. Lowry, Capt Ralph W. Measuring Productivity of Depot-Level Aircraft Maintenance in the Air Force Logistics Command. MS Thesis GSM/LSY/85S-21. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1985 (AD-A162308).

25. Montemayor, Jose O. Determining Optimal Resource Allocations Using Production Models Derived From Efficiency Analysis of Empirical Data. MS Thesis GLM/LSM/85S-53. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1985 (AD-A161082).

26. Neter, John and others. Applied Linear Regression Models. Homewood IL: Richard D. Irwin, Inc., 1983.

27. U.S. General Accounting Office. Report to the Chairman, Senate Committee on Government Affairs: Compendium of GAO's Views on the Cost Savings Proposals of the Grace Commission. Report GAO/OCG-85-1, Volume I. Washington: U.S. General Accounting Office, 19 February 1985.

28. U.S. General Accounting Office. The Department of Defense's Planning, Programming, and Budgeting System by the Joint DOD/GAO Working Group on PPBS. Report GAO/OACG-84-5. Washington: U.S. General Accounting Office, September 1983.

VITA

Captain Jay R. Wallace, II, was born on 24 September 1953 in Honolulu, Hawaii. He graduated from high school in Mascoutah, Illinois, in 1971 and attended the United States Air Force Academy from which he received the degree of Bachelor of Science in Management in June 1975. Upon completing pilot training in September 1976, he served as an F-4E pilot in the 33rd Tactical Fighter Wing, Eglin AFB, Florida. After changing career fields in 1979, he served as a budget officer at Hurlburt Field, Florida; Incirlik AB, Turkey; and Headquarters Strategic Air Command, Offutt AFB, Nebraska. In 1984 he was reassigned as a cost analyst at HQ SAC until entering the School of Logistics, Air Force Institute of Technology, in June 1985.

Military address: HQ MAC/ACCC
Scott AFB, IL 62225

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A174371

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): AFIT/GSM/LSY/86S-21	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: School of Systems and Logistics; 6b. OFFICE SYMBOL (If applicable): AFIT/LS	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): Air Force Institute of Technology, Wright-Patterson AFB, Ohio 45433-6583	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification): See Box 19	

12. PERSONAL AUTHOR(S): Jay R. Wallace, II, B.S., Captain USAF

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
MS Thesis	FROM ______ TO ______	1986 September	109

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.
05	01	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Mathematical Programming, Regression Analysis, Efficiency, Finance, Budgets

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Title: FEASIBILITY OF USING DATA ENVELOPMENT ANALYSIS AND REGRESSION ANALYSIS IN COMBINATION TO INDICATE EFFICIENT ALLOCATION OF OPERATIONS AND MAINTENANCE FUNDS

Approved for public release: IAW AFR 190-1. E. WOLAVER 29 Sept 86
Dean for Research and Professional Development
Air Force Institute of Technology (AFC)
Wright-Patterson AFB OH 45433

Thesis Chairman: William F. Bowlin, Major, USAF
Assistant Professor of Quantitative Methods

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
William F. Bowlin, Major, USAF	513-255-4845	AFIT/LSY

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

The present call for efficient use of national defense resources draws its urgency from the highest level of the American government. One problem with the Air Force's current method of allocating Operations and Maintenance (O&M) funds is that historical spending inefficiencies may get carried into the baseline for future funds allocations, thereby defeating this call for efficient use of resources.

This research furthers the exploration of ways to efficiently allocate resources and studies the feasibility of using a relatively new methodology developed by Charnes, Cooper, and Rhodes called Data Envelopment Analysis (DEA), in combination with regression analysis, as a method for indicating efficient allocation of Air Force O&M funds. Intended to illustrate O&M funds in general, this study uses a data base consisting of Fiscal Year (FY) 1985 Base Operations Support (BOS) obligations and activity measures from 25 Strategic Air Command (SAC) wings.

Research methodology consists of first, the derivation of a regression model expressing BOS obligations as a function of BOS activity measures; second, the development of an efficient data set by using information generated by DEA; and finally, the application of regression analysis to efficient data in order to predict efficient BOS obligations for the SAC wings involved.

Research results indicate that if the sampled BOS obligations and activities had been at theoretically efficient levels, then the cumulative FY 1985 BOS obligations would have been $9.7 million less than actual. The study also concludes that these procedures show promise for use in establishing a funds allocation baseline based on efficient use of funding.

This research concludes with recommendations for further study into the areas of developing appropriate output measures, using indicator variables in the regression model, breaking down efficiency measures into their technical and scale efficiency components, and dealing with collinearity among efficient data.

END

12-86

DTIC